The Philosophical Magazine

FIRST PUBLISHED IN 1798

A Journal of Theoretical Experimental and Applied Physics

Vol. 3 September 1958 No. 33
Eighth Series

£1 5s. 0d., plus postage
Annual Subscription £13 10s. 0d., payable in advance

Printed and Published by
TAYLOR & FRANCIS LTD
RED LION COURT, FLEET STREET, LONDON, E.C.4

THE PHILOSOPHICAL MAGAZINE

Authors wishing to submit papers for publication in the Journal should send manuscripts directly to the Publishers.

Manuscripts should be typed in *double* spacing on one side of quarto (8×10 in.) paper, and authors are urged to aim at absolute clarity of meaning and an attractive presentation of their texts.

References should be listed at the end in alphabetical order of authors and should be cited in the text in terms of author's name and date. Diagrams should normally be in Indian ink on white card, with lettering in soft pencil, the captions being typed on a separate sheet.

A leaflet giving detailed instructions to authors on the preparation of papers is available on request from the Publishers.

Authors are entitled to receive 25 offprints of a paper in the Journal free of charge, and additional offprints can be obtained from the Publishers.

The *Philosophical Magazine* and its companion journal, *Advances in Physics*, will accept papers for publication in experimental and theoretical physics. The *Philosophical Magazine* publishes contributions describing new results, letters to the editor and book reviews. *Advances in Physics* publishes articles surveying the present state of knowledge in any branch of the science in which recent progress has been made. The editors welcome contributions from overseas as well as from the United Kingdom, and papers may be published in English, French and German.

Random Walks and Drift in Chemical Diffusion†

By A. D. Le Claire
Metallurgy Division, Atomic Energy Research Establishment, Harwell, Berks.

[Received April 15, 1958]

Abstract

In self diffusion the average displacement $\overline{X}(t)$ of an atom after a time t is zero. It is pointed out that in chemical diffusion $\overline{X}$ is not necessarily zero, for there are several mechanisms tending to make an atom jump preferentially in one direction rather than the other along a chemical concentration gradient, i.e. tending to produce a drift of atoms superimposed upon their otherwise random movement. This leads to a substantial modification of the classical Einstein equation $D=\overline{X^2}(t)/2t$ connecting the self-diffusion coefficient D with the mean square displacement $\overline{X^2}(t)$ and the equation for chemical diffusion which replaces it is derived. This is shown to lead, on the basis of a simple model, to Darken's equation for the chemical diffusion coefficient.

Although in both this and a more general model there are usually two independent mechanisms contributing to $\overline{X}$, it turns out to be an interesting feature of chemical diffusion that only *one* of these is manifest in the chemical diffusion coefficient. The total drift $\overline{X}$ can be measured independently of the diffusion coefficient and values calculated from the theoretical expressions derived agree well with experimental measurements for zinc and for copper in α-brass/copper diffusion couples.

The results are generalized for chemical diffusion in multicomponent systems and expressions obtained for the constants in the familiar schemes of equations used to describe multicomponent diffusion. It is concluded that the cross terms L_{ij} in the Onsager scheme are *not* primarily the result of correlation between the directions of successive jumps of atoms, as has been suggested.

§ 1. Introduction

The atomic process responsible for the observed effects of diffusion is generally accepted to be the Brownian motion which each atom performs as a result of thermal agitation. In particular, in a solid crystal, each atom is visualized as spending only a finite time on any one lattice site, then jumping to a neighbouring site, then to another, and so on, in this way persuing an endless path or 'random walk' throughout the crystal. In substitutional solid solution the jumps are most probably made into vacant lattice sites (vacancy diffusion) and a particular atom only has an opportunity to jump when a vacant site becomes adjacent to it as a result of the movements of the other atoms. In interstitial solid solution the rate of jumping of solute atoms is not limited by the availability of vacant sites, at least in dilute solutions.

The only features of such a 'random walk' which are usually accessible to experiment are statistical quantities like $\bar{X}(t)$ and $\overline{X^2}(t)$, the average nett displacement and the average square of the nett displacement X of an atom

† Communicated by the Author.

after a time t, measured along a given direction, the x-axis, the averages being taken over a very large number of identical particles. For example, there exists the simple and well known Einstein (1905) relation between $\overline{X^2}(t)$ and the diffusion coefficient D, which is measured by determining the nett flux of atoms across a plane arising from their random movements when there exists a concentration gradient of the atoms

$$D = \tfrac{1}{2}\overline{X^2}/t. \qquad (1)$$

However, as we shall discuss shortly, this equation is valid only for random walks which take place in chemically homogeneous systems (Self Diffusion). It is the purpose of this paper to discuss the problem of random walks in systems with a chemical concentration gradient (Chemical Diffusion) and to derive the equation which replaces (1) in this case. Although the relation between chemical diffusion coefficients and self-diffusion coefficient has been discussed in the past, largely in thermodynamic terms (Darken 1948, Bardeen 1949, Bardeen and Herring 1951, Le Claire 1953), little attention seems to have been given to a treatment in terms of the statistical features of the associated random walk. We shall find that although the well known Darken equation follows again from such an approach there is the further interesting result that in a chemical concentration gradient atoms suffer a nett displacement $\bar{X}$, only part of which is made manifest in a chemical diffusion experiment.

We begin by outlining a derivation of eqn. (1) which is then adapted to the chemical diffusion case.

§ 2. Self Diffusion

Consider a crystal in which there exists along the x direction a concentration gradient of radioactive atoms of one species but which is otherwise homogeneous chemically. We wish to know the nett rate of transfer of atoms across some reference plane x_0, arising from their random motion.

Consider an infinitesimal layer dx_a at x_a. At $t=0$ each atom within this layer begins a random walk. Let $f(X, t)$ be the relative probability that in time t an atom will have migrated a distance X from x_a, measured along the x-axis. Since the medium is homogeneous chemically, migrations of $+X$ and $-X$ are equally probable, i.e. there is no tendency for an atom to drift preferentially in one direction or the other, so that the average displacement after t of a large number of atoms is zero.

$$\bar{X} = \int_{-\infty}^{+\infty} Xf(Xt)\,dX = 0. \qquad (2)$$

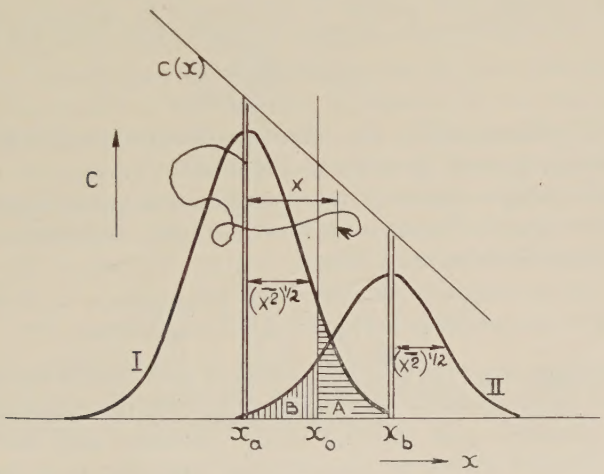
Similarly, all odd moments of $f(X, t)$, $\overline{X^3}$, $\overline{X^5}$ etc. are zero.

But the average squared displacement $\overline{X^2}$ is not zero and after t all the radioactive tracer atoms originally within the layer dx_a will be spread out and distributed in a manner indicated by curve I, symmetrical about x_a because $\bar{X}=0$. (The curve can readily be shown to be Gaussian, as drawn, but its precise shape does not concern us here.) The area under

the curve is $c(x_a)\,dx_a$, $c(x)$ being the tracer concentration at x, and being Gaussian its half width is $(\overline{X^2})^{1/2}$. A similar curve II represents the distribution at t of those tracer atoms originally contained in an infinitesimal layer at x_b an equal distance on the right-hand side of x_0. Its area is $c(x_b)\,dx_b$ but its width is the same as curve I.

The shaded area A represents the number of tracer atoms from x_a which are on the right-hand side of x_0 after t and area B the number from x_b now on the left-hand side of x_0. Because A > B there has been a nett flow of atoms across x_0 and in the direction of decreasing concentration: the total nett flow is to be got by summing the differences A − B for all pairs of regions like dx_a and dx_b.

Fig. 1



The total number of tracer atoms originally on the left-hand side of x_0 which will be found on the right-hand side after t is

$$\int_{-\infty}^{x_0} c(x) \left\{ \int_{x_0-x}^{\infty} f(Xt)\,dX \right\} dx. \qquad \ldots\ldots \quad (3)$$

The total integrand represents an area such as A. Similarly, the total number of atoms which have crossed x_0 from right to left, i.e. the sum of areas like B, is

$$\int_{x_0}^{\infty} c(x) \left\{ \int_{-\infty}^{x_0-x} f(Xt)\,dX \right\} dx. \qquad \ldots\ldots \quad (4)$$

The difference between (3) and (4) gives the nett transfer across x_0 from left to right. Expanding $c(x)$ about its value at x_0

$$c(x) = c(x_0) + \frac{\partial c}{\partial x}(x-x_0) + \frac{1}{2}\frac{\partial^2 c}{\partial x^2}(x-x_0)^2 + \ldots \qquad \ldots \quad (5)$$

and substituting in (3) and (4), the difference between them becomes

$$c(x_0)\left[\int_{-\infty}^{x_0}\left(\int_{x_0-x}^{\infty} f(Xt)\,dX\right)dx-\int_{x_0}^{+\infty}\left(\int_{-\infty}^{x_0-x} f(Xt)\,dX\right)dx\right]+\frac{\partial c}{\partial x}$$
$$\times\left[\int_{-\infty}^{x_0}(x-x_0)\left(\int_{x_0-x}^{\infty} f(Xt)\,dX\right)dx-\int_{x_0}^{+\infty}(x-x_0)\left(\int_{-\infty}^{x_0-x} f(Xt)\,dX\right)dx\right]$$
$$+\frac{1}{2}\frac{d^2c}{dx^2}\left[\int_{-\infty}^{x_0}(x-x_0)^2\left(\int_{x_0-x}^{\infty} f(Xt)\,dX\right)dx\right.$$
$$\left.-\int_{x_0}^{+\infty}(x-x_0)^2\left(\int_{-\infty}^{x_0-x} f(Xt)\,dX\right)dx\right]+\ldots. \qquad (6)$$

Integrating by parts twice and dividing by t to give the average nett rate of flow J, we find

$$J=\left(\frac{1}{t}\right)\left\{c(x_0)\bar{X}-\frac{\partial c}{\partial x}\left(\frac{\overline{X^2}}{2}\right)+\frac{\partial^2 c}{\partial x^2}\left(\frac{\overline{X^3}}{3!}\right)+\ldots\right\}. \qquad (7)$$

Since for self-diffusion $\bar{X}$, $\overline{X^3}$ etc. are zero, this becomes Fick's law

$$J=-\left(\frac{\overline{X^2}}{2t}\right)\frac{\partial c}{\partial x} \qquad (8)$$

and hence the relation (1) for D. We note that terms of order $\partial^3 c/\partial x^3$ and higher are being ignored in order to obtain D.

If Γ is the average number of jumps made by an atom in unit time, the total number of jumps after time t is $N=\Gamma t$. Let x_j denote the x component of the jth jump of an atom. Then

$$\overline{X^2}=\overline{\left(\sum_{j=1}^{N} x_j\right)^2}=\overline{\sum_{j=1}^{N} x_j^2}+\overline{\sum\sum_{j\neq j'} x_j x_{j'}}. \qquad (9)$$

Since the averages are over a large number of atom paths the last term will be zero, if we ignore possible correlations between the directions of successive jumps (Bardeen and Herring 1951, Compaan and Haven 1956, Le Claire and Lidiard 1956) for positive and negative products will occur with equal frequency. Suppose now that when an atom makes a jump it has a choice of s directions in which it can move. Let Γ_i be the rate at which atoms make jumps in the direction i, so that $\Gamma=\sum_{i=1}^{s}\Gamma_i$, and let x_i be the x component of the corresponding jump vector. Then

$$\overline{X^2}(t)=\overline{\sum_{j=1}^{N} x_j^2}=N\overline{x_j^2}=(N/\Gamma)\sum_{i=1}^{s}\Gamma_i x_i^2=t\sum\Gamma_i x_i^2 \qquad (10)$$

so that we have for D

$$D=\tfrac{1}{2}\sum_{i=1}^{s}\Gamma_i x_i^2. \qquad (11)$$

In self diffusion the Γ_i are all equal in cubic crystals but not necessarily so an anisotropic crystals. For cubic crystals we may therefore write

$$D=\tfrac{1}{2}\Gamma_i\sum_{i=1}^{s} x_i^2=\tfrac{1}{2}\Gamma\overline{x_j^2} \qquad (12)$$

where in the last expression $\overline{x_j^2}$ is now an unweighted mean.

Note added in proof.—'Random walk' treatments of self diffusion have also been given, among others, by Kramers (1940) and Chandrasekhar (1943).

§ 3. Chemical Diffusion

It is well established that both self-diffusion coefficients and chemical diffusion coefficients vary with concentration so that the atomic jump rates must change continuously as we move along a concentration gradient. This means that we can no longer assume, as we did above, that the probability function $f(X, t)$ is a function only of X and t. The probability that an atom migrates a distances X in time t will now depend also on the distribution of concentration throughout the whole region of the crystal which is accessible to the atom in time t.

Generally, during the course of diffusion, this distribution will change continuously so that a proper specification of f for all t would be very complex. We can however consider the special case of steady state diffusion, e.g. the steady flow of material into, through and out of a plate. Here, the distribution of concentration within the plate does not change with time when the steady state is reached.

Under these conditions the probability that in time t an atom starting from any given position will migrate a distance X, in addition to being a function of X and t, will depend also only on the position of the starting point and on the nature of the whole but steady distribution of concentration. The position of the starting point we can specify by the concentration there so that for a random walk in a steady concentration gradient we replace $f(X, t)$ by $f(X, t, c(x))$, which is the probability that an atom which starts from x where the chemical concentration is $c(x)$ will migrate X in time t. The precise form of f will depend on the nature of the stationary concentration distribution to which it refers.

$f(X, t, c(x))$ can be expanded about its value for atoms commencing their random walk at the reference plane x_0,

$$\begin{aligned} f(X, t, c(x)) = f(X, t, c(x_0)) \\ + \frac{\partial f}{\partial c}\left\{\frac{\partial c}{\partial x}(x-x_0) + \frac{1}{2}\frac{\partial^2 c}{\partial x^2}(x-x_0)^2 + \ldots\right\} \\ + \frac{1}{2}\frac{\partial^2 f}{\partial c^2}\left\{\left(\frac{\partial c}{\partial x}\right)^2 (x-x_0)^2 + \ldots\right\}. \quad \ldots \quad (13) \end{aligned}$$

We now substitute this expression for f into eqn. (6), perform twice an integration by parts and divide by t to give J, the average *rate* of flow across x_0. We find for J

$$\begin{aligned} J = c(x_0)(\bar{X}/t) - \frac{\partial c}{\partial x}\left(\frac{\overline{X^2}}{2t}\right) - c(x_0)\frac{\partial c}{\partial x}\left(\frac{\partial}{\partial c}\cdot\frac{\overline{X^2}}{2t}\right) \\ + \frac{1}{2}\frac{\partial^2 c}{\partial x^2}\left(\frac{\overline{X^3}}{3t}\right) + \left(\frac{\partial c}{\partial x}\right)^2\left(\frac{\partial}{\partial c}\frac{\overline{X^3}}{3t}\right) + \ldots. \quad \ldots \quad (14) \end{aligned}$$

In this expression the moments are to be evaluated for atom paths beginning at x_0.

Previously we knew from the symmetry of $f(Xt)$ that its odd moments X, $\overline{X^3}$ etc. would be zero. Whether or not this is true also for $f(X, t, c(x))$ depends on the detailed model of atomic jumps in a concentration gradient.

In general it will *not* be true. Therefore, since $\bar{X}$ is to a good approximation proportional to dc/dx (see later) we arrive at Fick's law in chemical diffusion, that J is proportional only to the concentration gradient, provided we can ignore as being negligible terms in $\partial^2 c/\partial x^2$, $(\partial c/\partial x)^2$ etc. In self-diffusion we were obliged to ignore only terms of higher order than these. Fick's law is therefore for chemical diffusion a grosser approximation than for self-diffusion and we might expect the law to break down first for diffusion in sufficiently steep chemical concentration gradients. It would be interesting to study this breakdown experimentally, perhaps by looking for a time dependence of the apparent diffusion coefficient during the initial stages of diffusion across an interface separating two materials the compositions of which differ as much as possible. It would be preferable to study the diffusion of an interstitial solute, for in substitutional solutions Kirkendall effect phenomena, particularly non-equilibrium vacancy concentrations, might well play a predominant part in producing a time dependent D (Seitz 1955).

In that follows we shall assume that the gradient is small so that all but the first three terms of (14) can be neglected: we can then define a chemical diffusion coefficient as

$$D = -c(x_0)(\bar{X}/t)(\partial c/\partial x)^{-1} + \overline{X^2}/2t + c(x_0)(\partial/\partial c)(\overline{X^2}/2t). \qquad (15)$$

This expression for D replaces for chemical diffusion the classical Einstein expression which is appropriate only for self-diffusion.

Equation (14) shows that in general the nett rate of flow across a given plane is characterized by a drift motion of the atoms, represented by the term in $\bar{X}$, superimposed on the otherwise random motion of the atoms, represented by the $\overline{X^2}$ term as in self-diffusion: there is an additional contribution arising from the change in the extent of the random motion with concentration and therefore with position—the third term in eqn. (14).

Further discussion of (14) or (15) must be based on particular models of the jump process and the dependence of its rate on concentration. For this purpose we shall need to evaluate $\bar{X}$ and $\overline{X^2}$.

$\bar{X}$ is given by

$$\bar{X} = \overline{\left(\sum_{j=1}^{N} x_j\right)} = N\bar{x}_j = t\sum_{i=1}^{s} \Gamma_i x_i. \qquad (16)$$

$\overline{X^2}$ is again given by (9) but the second term $\overline{\sum\sum_{j\neq j'} x_j x_{j'}}$ is no longer zero when atoms experience a drift, i.e. when they have a tendency to jump in one direction in preference to the opposite along the concentration gradient. It can be shown quite generally that, if N is large

$$\overline{X^2} = N\overline{x_j^2} + (\bar{X})^2. \qquad (17)$$

Consider as a simple example a one dimensional crystal with lattice site spacing a. Let p_+ and p_- be the relative probabilities that when a jump is made it is by $+a$ or $-a$ respectively. Then $\bar{X} = (p_+ - p_-)Na$.

There are $N(N-1) \sim N^2$ terms in the sum $\sum\sum x_j x_{j'}$ and each is of absolute value a^2. The probability that a term is positive is $p_+^2 + p_-^2$ and that it is negative is $2p_+p_-$. The summation term is therefore

$$\overline{\sum\sum x_j x_{j'}} = N^2(p_+^2 + p_-^2 - 2p_+p_-)a^2 = (\bar{X})^2$$

and hence equation (17).

However, in the example we discuss later $(\bar{X})^2$ is much less than $N\overline{x_j^2}$ ($\sim 4{\cdot}5\%$). We shall assume this is generally true and therefore ignore it, leaving eqn. (10) still valid for diffusion with drift. (The error involved is smaller than that arising from ignoring correlation between the directions of successive jumps, which is $\sim 20\%$.) Since it is proportional (v.i.) to $(dc/dx)^2$, a value of $(\bar{X})^2$ comparable with $\overline{X^2}$ would lead to a term in J proportional to $(dc/dx)^3$ and so contribute to the breakdown of Fick's law in chemical diffusion.

In arriving at the above equations, the averages are taken over a large number of atom paths each of which involves the same number N of jumps. Also, the set of values Γ_i (or p_+ and p_-) is assumed to be the same at each lattice point. Strictly, these conditions do not hold for a random walk in a chemical gradient. The Γ_i change continuously along the direction of the gradient and N, the number of jumps made in a given time, will vary with the direction of the path. However, for sufficiently small t, so that not too extensive ranges of values of Γ_i or of N are involved, we shall assume as a reasonable approximation for our present purpose that the true values of $\bar{X}$ and $\overline{X^2}$ do in fact approach those given by (16) and (10).

§ 4. Calculation of $\bar{X}$, $\overline{X^2}$ and D for a Simple Model

We shall first discuss a simple model which reveals the essential features of random walks in chemical diffusion and then generalize the results in a later section.

The central height of the potential barrier over which an atom jumps in moving from one site to another we assume to be determined by the average of the concentrations at these two sites†. This height will vary as we move along the concentration gradient and if this were the only factor influencing the random walk we could represent the situation by the potential barrier diagram of fig. 2. Between two sites, say 1 and 2, the rate of jumping Γ is the same for the direction $1 \rightarrow 2$ as for $2 \rightarrow 1$, and by definition is equal to the rate of jumping in a homogeneous alloy of composition corresponding to a point midway between 1 and 2, i.e. the self-diffusion jump rate.

Each atom will clearly experience a drift on account of the changing barrier height for an atom at a given site will always move with higher probability to the right than to the left in fig. 2.

† By " concentration at a site " is meant of course the average concentration on the lattice plane through that site and normal to the direction of the concentration gradient.

But increasing the concentration gradient while keeping constant the average concentration between two sites is unlikely to leave the jump rate between them unaffected and we therefore expect also some dependence of the jump rate on the differences in concentration at the two sites.

Fig. 2

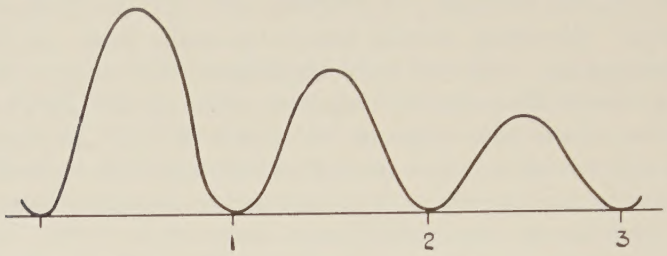
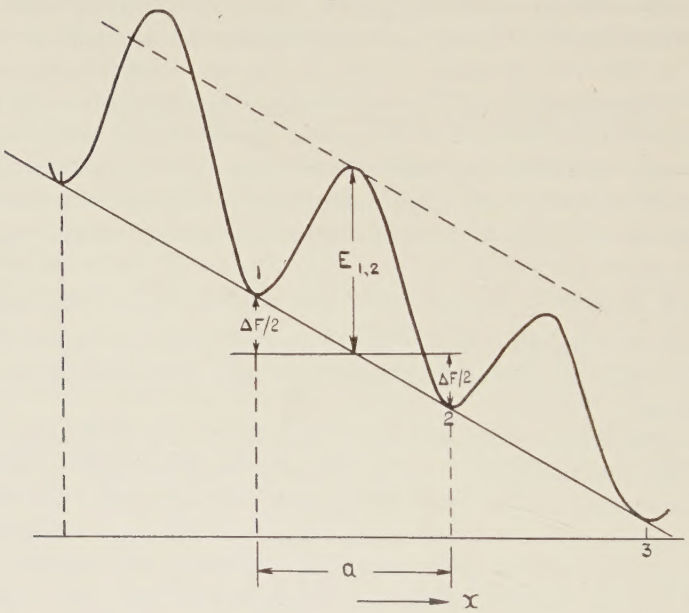


Fig. 3



Let us assume that the affect on Γ of a difference in concentration arises because the solid solution is not ideal so that there is a change ΔF in free energy when an atom moves from one site at one concentration to another at a different concentration. As a result of this the effective height of the potential barrier is increased by $\frac{1}{2}\Delta F$ for a jump in the direction which results in an increase in energy and is reduced by the same amount for the same jump in the opposite direction. The effect on the potential barrier diagram is as shown in fig. 3. The vertical height, e.g. $E_{1\cdot2}$ of each barrier is the same as in fig. 2 but the minimum on one side of each

barrier is raised by $\frac{1}{2}\Delta F$ and on the other side depressed by the same amount. A jump from $1 \rightarrow 2$ then produces a nett energy change ΔF.

This effect clearly provides an additional cause for drift since an atom can now jump even more readily to the right than to the left from any given site.

These two components of the drift $\bar{X}$, one due to the variation of jump rate with average composition, the other to its variation with difference in composition, can of course be oppositely directed. As drawn, fig. 3 provides for both components being in the same direction.

In calculating $\bar{X}$ and $\overline{X^2}$ we shall assume a vacancy mechanism of diffusion. There is then a possible third component of drift if there is a gradient of vacancy concentration for atoms will tend to move towards regions of high vacancy concentration. However we shall see that if the concentration of vacancies is everywhere that corresponding to equilibrium with the local chemical concentration this drift component is zero, as might be expected.

Let $\Gamma_{2\cdot3}$ be the rate of jumping of atoms from site 2 into a vacancy on site 3 and $N_{\mathrm{v}\cdot3}$ the fractional vacancy concentration on site 3. Then for a 2–3 jump Γ_i in eqns. (10) and (16) is $N_{\mathrm{v}\cdot3} \cdot \Gamma_{2\cdot3}$. If a denotes the interatomic spacing along the concentration gradient then for atoms starting at site 2

$$\begin{aligned}\bar{X}_2 &= t(aN_{\mathrm{v}\cdot3}\Gamma_{2\cdot3} - aN_{\mathrm{v}\cdot1}\Gamma_{2\cdot1}) \\ &= atN_{\mathrm{v}\cdot2}(\Gamma_{2\cdot3} - \Gamma_{2\cdot1}) + 2a^2\Gamma_2{}^*t\frac{\partial N_{\mathrm{v}}}{\partial x} \qquad \ldots \quad (18)\end{aligned}$$

where $\Gamma_2{}^*$ is the rate of jumping into vacancies in a homogeneous alloy of composition equal to that on site 2. Similarly, to a first order

$$\begin{aligned}\overline{X_2{}^2} &= t(a^2N_{\mathrm{v}\cdot3}\Gamma_{2\cdot3} + a^2N_{\mathrm{v}\cdot1}\Gamma_{2\cdot1}) \\ &= 2a^2tN_{\mathrm{v}\cdot2}\Gamma_2{}^*. \qquad \ldots \quad (19)\end{aligned}$$

Let $\mu'_{\mathrm{A}\cdot2}$ denote the free energy per atom of the diffusing species A at position 2, *less* the contribution due to entropy of mixing. $\mu'_{\mathrm{V}\cdot2}$ denotes the corresponding quantity for a vacancy at 2. When an atom jumps from $2 \rightarrow 3$ a vacancy moves simultaneously in the opposite direction so that the energy change ΔF is

$$\Delta F_{2\cdot3} = (\mu'_{\mathrm{A}\cdot3} - \mu'_{\mathrm{A}\cdot2}) + (\mu'_{\mathrm{V}\cdot2} - \mu'_{\mathrm{V}\cdot3}) = \left(\frac{\partial\mu'_{\mathrm{A}}}{\partial c} - \frac{\partial\mu'_{\mathrm{V}}}{\partial c}\right)\frac{\partial c}{\partial x}a = -\Delta F_{2\cdot1}.$$

If E denotes the central height of a barrier the total effective barrier heights for jump 2–3 and 2–1 are

$$\left.\begin{aligned}Q_{2\cdot3} &= E_{2\cdot3} + \frac{1}{2}\left\{\frac{\partial\mu'_{\mathrm{A}}}{\partial c} - \frac{\partial\mu'_{\mathrm{V}}}{\partial c}\right\}\frac{\partial c}{\partial x}a \\ Q_{2\cdot1} &= E_{2\cdot1} - \frac{1}{2}\left\{\frac{\partial\mu'_{\mathrm{A}}}{\partial c} - \frac{\partial\mu'_{\mathrm{V}}}{\partial c}\right\}\frac{\partial c}{\partial x}a.\end{aligned}\right\} \qquad \ldots \quad (20)$$

The rates of jumping $\Gamma_{2\cdot3}$ and $\Gamma_{2\cdot1}$ can be written

$$\Gamma_{2\cdot3}=\nu \exp(-Q_{2\cdot3}/kT) \text{ and } \Gamma_{2\cdot1}=\nu \exp(-Q_{2\cdot1}/kT)$$

ν being the vibration frequency of an atom at 2. Since $Q_{2\cdot3}-Q_{2\cdot1}$ will be very small we can write for the difference between the Γ's

$$\Gamma_{2\cdot3}-\Gamma_{2\cdot1}=-\Gamma_2{}^*\,.\,(Q_{2\cdot3}-Q_{2\cdot1})/kT. \quad . \quad . \quad . \quad . \quad (21)$$

Substituting (20) into (21) and then (21) into (18) and making use of the relation for the complete chemical potential of the vacancies, $\mu_V=\mu'_V+kT\ln N_V$, we find for $\bar{X}$

$$\bar{X}=\left\{\frac{\partial}{\partial c}(\tfrac{1}{2}\overline{X^2})-(\tfrac{1}{2}\overline{X^2})\frac{1}{kT}\frac{\partial\mu'_A}{\partial c}+(\tfrac{1}{2}\overline{X^2})\frac{1}{kT}\frac{\partial\mu_V}{\partial c}\right\}\frac{\partial c}{\partial x} \quad . \quad . \quad (22)$$

where $\overline{X^2}$ is

$$\overline{X^2}=2a^2tN_V\nu \exp(-E_2/kT) \quad . \quad . \quad . \quad . \quad . \quad (23)$$

and

$$\frac{\partial}{\partial c}(\tfrac{1}{2}\overline{X^2})=-\left\{\frac{\partial E}{\partial c}-\frac{kT}{N_V}\frac{\partial N_V}{\partial c}\right\}(\overline{X^2}/2kT). \quad . \quad . \quad . \quad (24)$$

The three terms in $\bar{X}$ correspond to the three drift components discussed above. The first derives from the variation of jump rate with average composition, the second from its dependence on the difference in composition and the third from the vacancy gradient. When (22) is put into (14), the equation for J, we find

$$J=-\frac{\overline{X^2}}{2t}\left\{1+\frac{c}{kT}\frac{\partial\mu'_A}{\partial c}-\frac{c}{kT}\frac{\partial\mu_V}{\partial c}\right\}\frac{\partial c}{\partial x}. \quad . \quad . \quad . \quad . \quad (25)$$

When the vacancies are in local thermodynamic equilibrium their gradient is such that μ_V is everywhere zero and all terms in μ_V disappear from the above equations. Also $\overline{X^2}$ can be replaced by $\overline{X^2}^*$, its value in self-diffusion. We then arrive for the chemical diffusion coefficient at the usual Darken equation

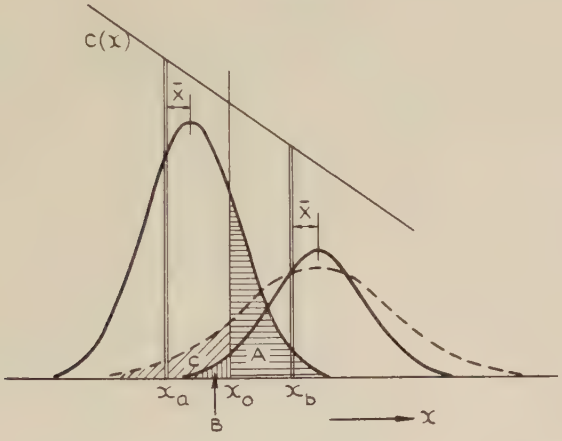
$$D=\frac{\overline{X^2}^*}{2t}\left\{1+\frac{\partial\ln\gamma_A}{\partial\ln c}\right\}=D^*\left\{1+\frac{\partial\ln\gamma_A}{\partial\ln c}\right\} \quad . \quad . \quad . \quad (26)$$

where γ_A is the activity coefficient, defined by $\mu'_A=kT\ln\gamma_A$.

It is to be noted that the first component of the drift, the term in $(\partial/\partial c)\overline{X^2}/2t$ has disappeared in the expressions for nett flux J and for D. That $\overline{X^2}$ varies with c produces on the present model both a contribution to the drift of atoms and a change in the rate of spreading out of atoms from a given plane as we move along the concentration gradient—the last term in the top line of (14). To the degree of approximation employed in the present calculation these two effects cancel out in J. This property is illustrated in fig. 4. On account of drift alone the two curves of fig. 1 are now each displaced by $\bar{X}$ from x_a and x_b respectively. There is an increase in J because the difference of areas A–B is now larger than before. But

because of the increase in rate of spreading as we move along the x-axis the curve about x_b should be wider than that about x_a and so is redrawn as a dashed curve. On account of this the nett flux J is reduced by the area C. The calculation shows that the area C is exactly equal to that part of the increase in A–B which is due to the $(\partial/\partial c)(\overline{X^2}/2t)$ component of drift. (The curves are shown as Gaussian for simplicity but in fact they will be skewed to the right. This does not affect the point of the argument.)

Fig. 4



The term $(D^*\partial \ln \gamma_A/\partial \ln c)$ in the Darken equation may then be interpreted as a drift component superimposed on the otherwise random motion of the atoms. It is important to appreciate that this is only *one* component of the total drift motion experienced by each atom. There are two components of drift, assuming the vacancies are in equilibrium, but only one is manifest in a measurement of a diffusion coefficient. Only when D^* is independent of concentration does $(D^*\partial \ln \gamma_A/\partial \ln c)$ represent the total drift.

Expressions can be readily derived for $\bar{X}$, $\overline{X^2}$ and $(\partial/\partial c)(\overline{X^2})$ for the vacancies themselves and these can be shown to combine in eqn. (14) to give for the nett flux of vacancies in a binary alloy of A and B the usual equation

$$J_v = (D_A - D_B)\frac{\partial N_A}{\partial x}. \qquad \ldots\ldots\ldots \quad (27)$$

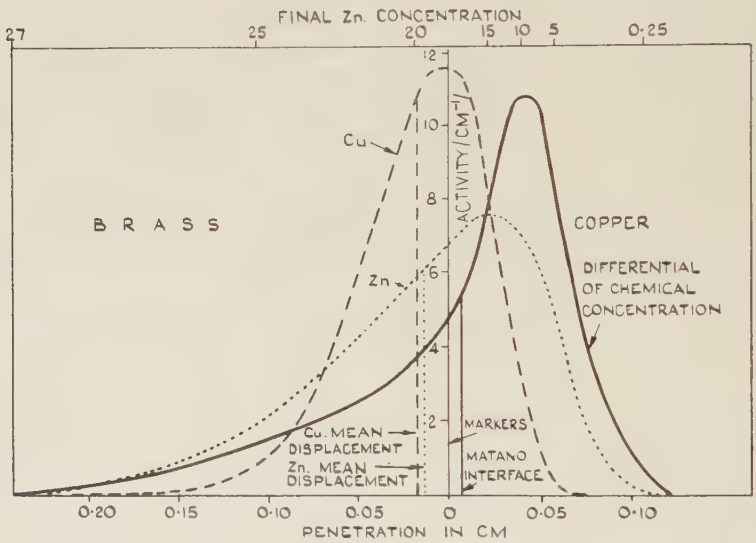
In arriving at (27) the vacancies are assumed to be everywhere in local equilibrium so that $\mu_V = 0$. In the full expression for J_v all terms derived from $\bar{X}_v^2$ and its variation with concentration disappear and J_v in eqn. (27) is constructed entirely from the remaining parts of $\bar{X}$ terms. The displacement of Kirkendall markers can then justifiably be described as due to a 'drift of vacancies'.

§ 5. Experimental Measurement of Drift of Atoms

A recent experiment by Shuttleworth† confirms that the drift suffered by individual atoms in a chemical concentration gradient is in fact fairly well represented by eqn. (22).

A thin foil of ~17% zinc α-brass was n-irradiated to produce copper and zinc, welded between a pair of discs, one of copper and one of 27% α-brass and then annealed to diffuse. The distribution of radioactive copper and zinc and of total zinc were then determined by counting and analysing sections sliced from the sample. The results are shown in fig. 5.

Fig. 5



The dotted line represented the distribution of radioactive zinc and the dashed line that of radioactive copper. The first moments of these distributions about the position of the marker interface give $\bar{\bar{X}}_{Cu}$ and $\bar{\bar{X}}_{Zn}$, the average displacements in the lattice of radioactive copper and zinc atoms from their original positions (the double bar denotes an average of $\bar{X}$ over the continuously changing concentration gradient necessarily experienced by atoms in an experiment of this type). Thus e.g.

$$\bar{\bar{X}}_{Cu} = \frac{\int c_{Cu^*}(x) x \, dx}{\int c_{Cu^*}(x) \, dx}. \qquad \ldots\ldots\ldots \quad (28)$$

The vertical dotted line in fig. 5 is drawn at a distance $\bar{\bar{X}}_{Zn}$ and the dashed line at a distance $\bar{\bar{X}}_{Cu}$ from the markers.

† Camagni, Pavollini and Shuttleworth, (to be published.) I am indebted to Dr. Shuttleworth for his kind permission to make use of his results prior to publication.

We note that the drift displacements are roughly of equal magnitude and both in the same direction, down the copper concentration gradient. This result is incompatible with any interpretation based on Darken's equation alone. Darken's equation reveals a drift flux of amount

$$-D^* \frac{\partial \ln \gamma}{\partial \ln c} \frac{\partial c}{\partial x}$$

and therefore a drift displacement of

$$\bar{X} = -D^* t \frac{\partial \ln \gamma}{\partial c} \cdot \frac{\partial c}{\partial x} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

in time t. $D^* \partial \ln \gamma \, \partial c$ for zinc is about ten times larger than it is for copper but in both cases is positive. Therefore, since $dc_{Zn} = -dc_{Cu}$, we should expect the drift for copper to be down the copper concentration gradient and that for zinc to be ten times greater and in the opposite direction—down the zinc concentration gradient. This being contrary to experiment confirms that that component of drift evident in Darken's equation cannot be the only one. We turn then to eqn. (22).

The last term of (22) we shall ignore for, if N_{V_0} is the equilibrium vacancy concentration,

$$\frac{1}{kT} \frac{\partial \mu_V}{\partial c} = \frac{\partial \ln (N_V / N_{V_0})}{\partial c} \sim \frac{N_V - N_{V_0}}{0{\cdot}3 N_{V_0}}$$

and this last quantity must be of order unity to compare even with the smaller of the two other terms of (22) (see the table). This requires that the vacancy concentration be $\sim 30\%$ above equilibrium, which is much greater than estimates ($\sim 1\%$ or less) of the excess vacancy concentration in couples of this type (Balluffi 1954, Barnes and Mazey 1958).

Let us now express $\bar{\bar{X}}_{Zn}$ and $\bar{\bar{X}}_{Cu}$ as

$$\bar{\bar{X}}_{Zn} = \overline{D^*_{Zn} t} \left\{ \frac{\overline{\partial \ln D_{Zn}{}^*}}{\partial c_{Zn}} - \frac{\overline{\partial \ln \gamma_{Zn}}}{\partial c_{Zn}} \right\} \overline{\frac{\partial c_{Zn}}{\partial x}} \quad . \quad . \quad (30\,a)$$

$$\bar{\bar{X}}_{Cu} = \overline{D^*_{Cu} t} \left\{ \frac{\overline{\partial \ln D_{Cu}{}^*}}{\partial c_{Cu}} - \frac{\overline{\partial \ln \gamma_{Cu}}}{\partial c_{Cu}} \right\} \overline{\frac{\partial c_{Cu}}{\partial x}} . \quad . \quad (30\,b)$$

The first three terms on the right-hand side of each equation are averages over the whole range of concentration and $\partial c / \partial x$ is to be averaged over the concentration range and over the time. The precise determination of these average values is a difficult problem in concentration dependent diffusion theory so that we shall be content with approximate estimates.

For $\overline{D^* t}$ we take simply one half of the second moments of the copper and zinc distributions in fig. 5 viz. $\frac{1}{2}\overline{X_{Cu}{}^2}$ and $\frac{1}{2}\overline{X_{Zn}{}^2}$, evaluated about the marker interface.

Diffusion coefficients in the α-brass system generally vary approximately logarithmically with composition so that $\partial \ln D^* / \partial c$ is fairly constant over the diffusion zone. Its values for copper and zinc have been estimated from the data of Inman *et al.* (1954) and Hino *et al.* (1957).

$\partial \ln \gamma/\partial c$ was estimated at 10, 20 and 25 atm% zinc from the data of Herbenar, Siebert and Duffendack as smoothed by Lumsden (1952).

The average value of $\partial c/\partial x$ across the diffusion zone at the end of the experiment was estimated from the full curve in fig. 5 to be ~ 75 atm%/cm. Assuming c to be a function of x and t in the combination $x/t^{1/2}$, the time average of dc/dx at any given concentration is twice the final value of dc/dx. The space and time average of dc/dx we therefore take as 150 atm%/cm or 1·5 atm fraction/cm.

The table shows the values of these several quantities, together with the experimental and calculated values of $\bar{\bar{X}}_{Zn}$ and $\bar{\bar{X}}_{Cu}$ the latter being calculated for the 20 atm% value of $\partial \ln \gamma/\partial c$. The agreement between the theoretical and experimental values shown in the last two columns is as good as can be expected in view of the approximations made—perhaps surprisingly so†. In particular the theoretical drifts are both in the same direction and approximately equal. The experimental result that $\bar{\bar{X}}_{Cu} > \bar{\bar{X}}_{Zn}$ is not reproduced theoretically but since the difference is in any case small some refinement of the theory would be necessary to achieve this. Taking into consideration the possible third component of drift, due to a non-equilibrium vacancy distribution, does not help here.

	$\overline{\partial c/\partial x}$ atm fract/cm	$\partial \ln \gamma/\partial c$				$\frac{\partial \ln D^*}{\partial c}$	$\frac{1}{2}\frac{\bar{\bar{X}}^2}{D^* t}=$	$\bar{\bar{X}}$ Theor. (cm)	$\bar{\bar{X}}$ Exptl. (cm)
		10%	20%	25%	Zn.				
Zn	−1·5	3·6	4·6	5·6		10·5	0·0018	−0·016	−0·014
Cu	+1·5	0·40	1·15	1·7		−10·1	0·0008	−0·014	−0·017

This drift component would be directed towards the zinc-rich side of the sample ($-ve\ x$) for both zinc and copper and being proportional to D its effect would be only to enhance the difference between the theoretical values of $\bar{\bar{X}}_{Zn}$ and $\bar{\bar{X}}_{Cu}$ without changing their order of size. This vindicates the neglect of this term.

It appears then that that component of the drift which is due to the variation of D^* with composition is for both zinc and copper greater than the other component due to the solution being non-ideal, and therefore dominates the drift behaviour in this system.

The drift displacements can also be compared with the Kirkendall marker shift which was measured in the same experiment. The marker shift is given by

$$x_m = -(D_{Zn}^* - D_{Cu}^*)\left\{1 + \frac{\partial \ln \gamma}{\partial \ln c}\right\} 2t \left(\frac{\partial c}{\partial x}\right)_m$$

† Dr. Manning (University of Illinois) has more recently measured the drifts of silver and of cadmium in a number of α-AgCd alloys and his results provide further confirmation of the present theory (private communication).

where the D's and $(\partial c/\partial x)_m$ are to be evaluated at the marker composition. $(\partial c/\partial x)_m \sim 120$ atm% cm^{-1} and $D_{\mathrm{Zn}}^* \simeq 3D_{\mathrm{Cu}}^*$ so that if we assume $D_{\mathrm{Zn}}^* t = \frac{1}{2}\overline{X_{\mathrm{Zn}}^2}$

$$\frac{x_m}{\overline{X}_{\mathrm{Zn}}} = -\frac{(4/3)\left\{1+\dfrac{\partial \ln \gamma}{\partial \ln c}\right\}\left(\dfrac{\partial c}{\partial x}\right)}{\left\{\dfrac{\partial \ln D_{\mathrm{Zn}}}{\partial c_{\mathrm{Zn}}} - \dfrac{\partial \ln \gamma_{\mathrm{Zn}}}{\partial c_{\mathrm{Zn}}}\right\}} = 0{\cdot}35.$$

Similarly

$$\frac{x_m}{\overline{X}_{\mathrm{Cu}}} = 0{\cdot}58.$$

These estimates are in reasonable agreement with the experimental value of about one half.

It will be clear that the simplest type of drift experiment to interpret theoretically would be one in which the drift is allowed to take place in a steady concentration gradient, i.e. under conditions of steady state diffusion. One might establish a steady rate of flow of solute into, through and out of a solvent plate, then cut the plate in two normal to the direction of diffusion, deposit a thin layer of radioactive solute on one of the cut surfaces, reweld the two surfaces together, with markers, and continue the steady state diffusion anneal. Such an experiment could be carried out either for a substitutional or an interstitial alloy system.

§ 6. General Discussion

We have seen that chemical diffusion differs from self-diffusion in that superimposed upon the otherwise random spreading out of atoms from any given region there is a tendency for atoms to drift in one direction or the other, i.e. the expression for J contains both $\bar{X}$ and $\overline{X^2}$ terms. There are two (or three where vacancies are not in equilibrium) components to $\bar{X}$ but only one of these is manifest in a diffusion coefficient measurement. This latter result is not peculiar to the model we have chosen but is true at least for any model in which the rate of jumping between a pair of sites is a function only of the composition at those two sites.

Denote the composition at a site from which a jump is made by α and that at the site to which it is made by ω. Then we suppose the rate of jumping to be a function $\Gamma(\alpha, \omega)$. Any such function can always be expressed as a function of the sum and of the difference of the arguments

$$\Gamma(\alpha, \omega) \equiv \Gamma'(\alpha+\omega, \omega-\alpha) \quad . \quad . \quad . \quad . \quad . \quad (31)$$

and

$$\frac{\partial \Gamma}{\partial \alpha} = \frac{\partial \Gamma'}{\partial(\alpha+\omega)} - \frac{\partial \Gamma'}{\partial(\omega-\alpha)}; \quad \frac{\partial \Gamma}{\partial \omega} = \frac{\partial \Gamma'}{\partial(\alpha+\omega)} + \frac{\partial \Gamma'}{\partial(\omega-\alpha)}. \quad . \quad (32)$$

The average displacement from a site on the plane at $x=0$ is then from (16)

$$\bar{X} = t\sum_{i=1}^{s} x_i \Gamma_i(c(0), c(x_i)) = t\sum_i x_i \left\{\Gamma(c(0), c(0)) + \frac{\partial \Gamma_i}{\partial \omega}\frac{\partial c}{\partial x} x_i\right\}$$

$$= t\left(\sum_i x_i^2 \frac{\partial \Gamma_i}{\partial \omega}\right)\frac{\partial c}{\partial x} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

for $\Gamma(c(0)c(0))$ is the rate of jumping in a homogeneous alloy of composition $c(0)$ so that $\sum x_i\Gamma_i(c(0), c(0)) = 0$. Using (32), (33) can be written

$$\bar{X} = t\sum_i \left\{ x_i^2 \frac{\partial \Gamma_i'}{\partial(\alpha+\omega)} + x_i^2 \frac{\partial \Gamma_i'}{\partial(\omega-\alpha)} \right\} \frac{\partial c}{\partial x}. \quad . \quad . \quad . \quad (34)$$

$\bar{X}$ is the sum of two terms, one depending on the variation of jump rate with average composition, the other on its dependence on difference in composition.

Similarly

$$\overline{X^2} = t\sum_i x_i^2 \Gamma_i(c(0), c(x_i)) = t\sum x_i^2 \Gamma(c(0), c(0)) = \overline{X^2}^* \text{ for } \sum x_i^3 \partial\Gamma_i/\partial\omega = 0. \quad (35)$$

Then

$$\partial \overline{X^2}_i / \partial c = t\sum_i x_i^2 \left\{ \frac{\partial \Gamma_i}{\partial \alpha} + \frac{\partial \Gamma_i}{\partial \omega} \right\}$$

or, using (32)

$$\frac{\partial}{\partial c}(\tfrac{1}{2}\overline{X^2}) = t\sum_i x_i^2 \frac{\partial \Gamma_i'}{\partial(\omega+\alpha)}. \quad . \quad . \quad . \quad . \quad . \quad (36)$$

Inserting (34), (35), and (36) into (15),

$$D = D^* - c(0)\sum x_i^2 \frac{\partial \Gamma_i'}{\partial(\omega-\alpha)} \quad . \quad . \quad . \quad . \quad . \quad (37)$$

from which the first component of the drift in (34) is absent.

In many crystals the Γ_i are all equal so that $\partial\Gamma_i/\partial(\omega - \alpha)$ can be taken outside the summation to obtain

$$D = D^*\left(1 - 2c\frac{\partial \ln \Gamma_i'}{\partial(\omega-\alpha)}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

which is a generalized form of Darken's equation.

An assumption implicit in the treatment so far is that the random walk behaviour of an atom of a given species in an alloy is determined entirely by the concentration of that species alone, for the probability functions $f(X, t, c(x))$ and the jump rates $\Gamma(x, \omega)$ contain reference only to the concentration of the diffusing species. This is clearly inadequate in systems of three or more diffusing components in substitutional solid solution (or of two diffusing components in interstitial solution). The probability that an atom of a particular species k migrates a distance X in time t will in general depend on the concentration distribution of each species present and for a given distribution can be written $f(X, t, c_1(x) \ldots c_n(x))$ where $c_i(x)$ denotes the concentration of i at the starting point x and n is the number of components whose concentration must be specified fully to describe the composition. Expanding this function about its value for atoms starting their random walk at the reference plane x_0, to give an equation analogous with (13), substituting the expansion into the general eqn. (6), and integrating, we obtain for the rate of flow of the kth component

$$J_k = c_k(x_0)\left\{ \frac{\overline{X_k}}{t} - \sum_{j=1}^{n} \frac{\partial c_j}{\partial x}\frac{\partial}{\partial c_j}\left(\frac{\overline{X_k}^2}{2t}\right)\right\} - \frac{\partial c_k}{\partial x}\frac{\overline{X_k}^2}{2t} + \ldots \quad (39)$$

which is the generalized form of eqn. (14).

If now it is further assumed that the rate of jumping of species k between two sites is a function of the concentration on these two sites of the n species, $\Gamma(\alpha, \omega)$ is replaced by

$$\Gamma_k(\alpha_1, \alpha_2 \ldots \alpha_n; \omega_1, \omega_2 \ldots \omega_n) \equiv \Gamma'(\alpha_1+\omega_1, \ldots \alpha_n+\omega_n; \omega_1-\alpha_1, \ldots \omega_n-\alpha_n)$$

and eqn. (32) by a similar set, n in number for each k.

$\bar{X}_k$ and $\overline{X_k^2}$ can then be evaluated as before to lead eventually to the familiar set of equations for describing multicomponent diffusion

$$J_k = -\sum_j^n D_{kj} \frac{\partial c_j}{\partial x} \quad \ldots\ldots \quad (40)$$

where, if the $\Gamma_{i,k}$ for any given k are all equal,

$$D_{kk} = D_k^* \left(1 - 2c_k \frac{\partial \ln \Gamma_{i,k}}{\partial(\omega_k - \alpha_k)}\right); \quad \ldots\ldots \quad (41)$$

and

$$D_{kj} = -2c_k D_k^* \frac{\partial \ln \Gamma_{i,k}}{\partial(\omega_j - \alpha_j)}; \quad k \neq j \quad \ldots\ldots \quad (42)$$

The diagonal terms D_{kk} are identical in form with (37) while the cross terms D_{kj} are drift terms determined by the dependence of the jump rate of species k on the difference in composition at the two sites of component j. A measurement of the drift of species k would reveal a second component

$$\sum_j \frac{\partial c_j}{\partial x} \sum_i x_i^2 (\partial \Gamma_{i,k} / \partial(\omega_j + \alpha_j))$$

but this again does not contribute to the *nett* rate of flow.

The dependence of Γ_i on the difference in composition at two sites we can again associate with the change in free energy when an atom jumps. $\bar{X}$ then becomes for the kth component,

$$\overline{X_k} = \sum_j \left\{ \frac{\partial}{\partial c_j} \left(\frac{\overline{X_k^2}}{2} \right) - \frac{X_k^2}{2kT} \cdot \frac{\partial \mu_k'}{\partial c_j} \right\} \frac{\partial c_j}{\partial x} \quad \ldots\ldots \quad (43)$$

in place of (22). μ_k' is the chemical potential of k, less the contribution from entropy of mixing. J_k is therefore

$$J_k = -c_k \frac{\overline{X_k^2}}{2t} \cdot \frac{1}{kT} \sum_{j=1}^n \frac{\partial \mu_k'}{\partial c_j} \frac{\partial c_j}{\partial x} - \frac{\overline{X_k^2}}{2t} \frac{\partial c_k}{\partial x}. \quad \ldots \quad (44)$$

Since $\mu_k' = \mu_k - kT \ln c_k$, μ_k being the complete chemical potential of k, the last equation can be written like (40) with

$$D_{kj} = \frac{\overline{X_k^2}}{2t} \cdot \frac{c_k}{kT} \cdot \frac{\partial \mu_k}{\partial c_j} \quad \ldots\ldots \quad (45)$$

or simply as

$$J_k = -\frac{\overline{X_k^2}}{2t} \cdot \frac{c_k}{kT} \cdot \frac{\partial \mu_k}{\partial x}. \quad \ldots\ldots \quad (46)$$

An alternative description of multicomponent diffusion, obtained from consideration of the thermodynamics of irreversible processes, leads to the equations (see e.g. de Groot 1951).

$$J_k = \sum_j L_{kj} \frac{\partial \mu_j}{\partial x} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (47)$$

Comparing (47) with (46) we see that $L_{kk} = (\overline{X_k^2}/2t)(c_k/kT)$ but that the cross terms L_{kj} relating the rate of flow of a species k to the chemical potential gradient of a species j, are all zero. The model we have used is not then sufficiently detailed to reveal any effects represented by the cross terms in (47).

According to Bardeen and Herring (1951) cross-terms appear only when 'correlation' effects are taken into account; these have so far been neglected in our discussion. From the nature of vacancy diffusion (although not peculiar to it) the direction of one jump of an atom is not independent of the direction of its previous or earlier jumps, for the vacancy which effects one jump is available to effect the second jump in the opposite rather than in a random direction. Successive jumps are said to be 'correlated' with the result that the summation term $\sum\sum x_j x_{j'}$ in eqn. (9) for $\overline{X^2}$ is not zero, as we assumed it to be, even for self-diffusion. Calculations show that for self-diffusion the true value of $\overline{X^2}$ is the right-hand side of (10) multiplied by a geometrical factor of roughly 0·8, the 'correlation factor'.

In chemical diffusion $\sum\sum x_j x_{j'}$ is in any event not zero but correlation will contribute further to it and in amount of the same order of magnitude as in self-diffusion. The important point is however that taking correlation into account will only modify the value of $\overline{X_k^2}$ that is to be inserted into the general eqn. (39) for J_k. Proceeding then from (39) to (46) the correlation factor will be carried through to appear only as a modification of the diagonal terms L_{kk}. It can in no way appear *wholly* in cross-terms. However, the detailed calculation of $\overline{X_k^2}$ for chemical diffusion would no doubt reveal in $\sum\sum x_j x_{j'}$ terms proportional to the concentration or chemical potential gradients of species other than k. These terms could be separated out from L_{kk} to yield cross-terms but they would be of second order because they would carry with them the coefficient $\partial \mu_k/\partial x$: this would furthermore be an unnecessary complication of the equations. If cross terms are extracted in this way then it appears, at least with the present model, that they only represent the effect on correlation of introducing the chemical concentration gradient: the diagonal terms always contain the main effect, determined by the relative concentration and jump rates of the several species present, and in the simplest representation, without cross terms, they contain the whole effect.

This conclusion is contrary to that of Bardeen and Herring who considered that the whole effect of correlation was contained in the cross terms.

There may well of course be other contributions to the L_{kj} which only a more detailed model would reveal.

Acknowledgments

I am indebted to Dr. H. M. Finniston for his interest in this work and to Drs. A. J. Mortlock, G. V. Kidson and W. M. Lomer for many helpful comments during the preparation of the manuscript.

References

Balluffi, R., 1954, *Acta Met.*, **2,** 194.
Bardeen, J., 1949, *Phys. Rev.*, **76,** 1403.
Bardeen, J., and Herring, C., 1951, *Atom Movements* (American Society for Metals).
Barnes, R. S., and Mazey, D. J., 1958, *Acta Met.*, **6,** 1.
Chandrasekhar, S., 1943, *Rev. mod. Phys.*, **15,** 1.
Compaan, K., and Haven, Y., 1956, *Trans. Faraday Soc.*, **52,** 786.
Darken, L. S., 1948, *Trans. Amer. Inst. min. (metall.) Engrs*, **175,** 184.
Einstein, A., 1905, *Ann. Phys.*, **17,** 549.
de Groot, S. R., 1951, *Thermodynamics of Irreversible Process* (Amsterdam : North Holland Publishing Co.).
Hino, J., Tomizuka, C., and Wert, C., 1957, *Acta Met.*, **5,** 41.
Inman, M. R. C., Johnston, D., Mercer, W. L., and Shuttleworth, R., 1954, *Oxford Radioisotopes Conference*, vol. II, p. 85.
Kramers, H. A., 1940, *Physica*, **7,** 284.
Le Claire, A. D., 1953, *Progress in Metal Physics*, vol. IV (London: Pergamon Press), p. 245.
Le Claire, A. D., and Lidiard, A. B., 1956, *Phil. Mag.*, **1,** 518.
Lumsden, J., 1952, *Thermodynamics of Alloys* (London : Institute of Metals).
Seitz, F., 1955, *J. phys. Soc. Japan*, **10,** 679.

The Diffusion Constant, Mobility and Lifetime of Minority Carriers in Germanium containing Parallel Arrays of Dislocations†

By J. B. Arthur, A. F. Gibson, J. W. Granville and E. G. S. Paige
Royal Radar Establishment, Malvern

[Received April 28, 1958]

Abstract

The presence of a high density of parallel edge dislocations in N-type germanium is found to significantly enhance the diffusion of holes in a direction parallel to the dislocations. The apparent diffusion constant is therefore anisotropic. In P-type germanium on the other hand the diffusion constant is isotropic and the carrier lifetime anisotropic.

At high electric fields the drift mobility of holes in N-type germanium is found to be anisotropic with respect to the dislocation array, no comparable effect occurring in P-type material.

These results can be explained by a model which assumes that dislocations introduce an additional acceptor level approximately intermediate in energy between the conduction and valence energy bands.

§ 1. Introduction

In a recent paper Bell and Hogarth (1957) have shown that the diffusion length of minority carriers in germanium and silicon, measured by the travelling light spot technique, may be anisotropic if the crystal contains a high density of edge dislocations parallel to one another. The diffusion length measured parallel to the dislocation array was typically a factor of 2 or 3 greater than that measured perpendicular to the array. In addition Bell and Hogarth observed that the decay of photoconductivity in filaments or rods cut parallel and across the dislocations was significantly different, the latter having the lower apparent lifetime. The authors' demonstrated, however, that the latter was a surface effect.

In interpreting the anisotropy of diffusion length, Bell and Hogarth assumed that the diffusion constant, D, was isotropic and equal to the normal value given by the Einstein relation :

$$D = \frac{\mu kT}{q} \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

where μ is the drift mobility of the minority carriers, k Boltzmann's constant, T the absolute temperature and q the electronic charge. All the anisotropy observed, therefore, resided in the deduced carrier lifetime, τ.

It is well known that dislocations in germanium or silicon crystals act as efficient recombination centres (Wertheim and Pearson 1957). In the

† Communicated by the Authors.

travelling light spot experiment, however, the total path length of a carrier is of the order of 10^4 times greater than the diffusion length which, in the crystals used, is itself an order of magnitude greater than the average separation between dislocations. Hence the probability of a carrier encountering a dislocation is virtually independent of its point of origin. Bell and Hogarth pointed out that significant anisotropy could be obtained on the above model if two additional assumptions were made, namely :

(*a*) The dislocations were largely polygonized into ' walls ' ; and

(*b*) The dislocations were surrounded by a potential barrier which tended to exclude minority carriers from the high recombination region.

Assumption (*a*) essentially reduces the model to a two dimensional one and ensures that carriers diffusing at right angles to the dislocation array must cross at least some dislocations. Assumption (*b*) restricts the random walk of carriers diffusing parallel to the dislocations and ensures that few, if any, of them cross a dislocation. On this model any anisotropy of τ from 1 to ∞ may be obtained, depending on the degree of polygonization and the effective barrier height.

The object of the work to be described in the present paper was, primarily, to check the validity of the assumption that the diffusion constant, D, is isotropic and equal to the normal value in heavily dislocated material. Two approaches to the measurement of D have been made, namely drift mobility measurements and the simultaneous measurement of phase and amplitude in the travelling light spot experiment. We shall show that in N-type germanium D is not isotropic and the results allow a more complete model to be built up. In P-type germanium, on the other hand, D is isotropic and all the anisotropy appears to reside in τ. In addition measurements of drift mobility at high electric fields will be described, the results obtained giving valuable quantitative support for the model proposed.

§ 2. The Measurement of Hole Mobility and Diffusion Constant by the Drift Method

2.1. *Drift Mobility : Experimental*

A conventional drift mobility equipment was used (Haynes and Shockley 1951), the pulsed field, pulsed emitter arrangement being adopted. Filaments were cut from N-type germanium crystals with ' grown in ' edge dislocations (Bell and Hogarth 1957), the major filament axis lying parallel or normal to the dislocation array. Emitter and collector contacts, end connections and voltage probes were attached by conventional alloying techniques.

The relatively low minority carrier lifetime of the dislocated material, together with the non-uniformity of the dislocation density, markedly reduced the accuracy and reproducibility that can normally be obtained in this experiment. In addition it was necessary to use relatively high

resistivity material (~ 30 ohm cm) to maintain the lifetime at an acceptable value, which increased errors due to conductivity modulation. To minimize the latter effect measurements were made at various emitter currents and the results extrapolated to zero emitter current.

The diffusion constant of the carriers was deduced from measurements of the width of the carrier arrival pulse at the collector (Lawrance and Gibson 1952). The above sources of error, particularly conductivity modulation, are about an order of magnitude more important in the determination of D than in μ.

2.2. *Drift Mobility : Results*

The results obtained may be summarized as follows :

(1) The drift mobility parallel to and across the dislocation array was the same in all pairs of filaments and equal to the mobility in normal undislocated, material of the same resistivity.

(2) In some pairs of filaments the diffusion constant, deduced from the width of the hole pulse at the collector, was significantly greater than the normal value parallel to the dislocations but approximately equal to the normal value across the dislocations. Unfortunately a relatively long carrier lifetime is required if an accurate value of D is to be obtained by this technique. A long lifetime implies a low dislocation density and hence small anisotropy, so an uneasy compromise has to be reached. Illustrative of the results obtained from such a compromise are those given in table 1 for an N-type germanium crystal (designated PH34), which was also examined by Bell and Hogarth (1957).

Table 1

	Filament cut parallel to dislocations	Filament cut normal to dislocations
Mobility	$\mu_{\parallel} = 1850 \pm 100$ cm^2 volt^{-1} sec^{-1}	$\mu_{\perp} = 1950 \pm 100$ cm^2 volt^{-1} sec^{-1}
Diffusion constant	$D_{\parallel} = 400 \pm 100$ cm^2 sec^{-1}	$D_{\perp} = 80 \pm 30$ cm^2 sec^{-1}

Crystal PH34. Resistivity ~ 30 ohm cm.

It may be shown, by general arguments, that any disturbance (e.g. non-uniformity of filament cross section or resistivity) will increase the width of the collector pulse and hence the apparent value of D. It seems likely that the moderately high value of $D_{\perp}$ is due to non-uniformity or similar effects. The very large value of $D_{\parallel}$ cannot be explained so easily, however, and suggests that the dislocations materially assist the rate of diffusion of holes. In view, however, of the difficult compromise on which this experiment is based the results cannot be considered conclusive without additional evidence from the travelling light spot experiment.

§ 3. The Measurement of Diffusion Constant and Lifetime by the Travelling Light Spot Method

3.1. *The Travelling Light Spot : Experimental*

It has been shown by Avery and Gunn (1955) that if the phase and amplitude of the signal in a travelling light spot experiment are measured simultaneously the values of D and τ may be obtained uniquely. The phase shift, which is a measure of the transit time of carriers from the light spot to the detector contact, naturally increases with the light interruption frequency. Alternatively, it is, in principle, sufficient to measure the amplitude only at 2 or more interruption frequencies.

In the experiments to be described measurements have been made of amplitude and phase at 4 kc/s and amplitude only at 800 c/s. In this way three equations are obtained for two unknowns (D and τ) which allows a check on consistency to be made. With low lifetime dislocated material phase measurements at the lower frequency are of inadequate accuracy though they are used, with undislocated crystals, for periodic checks on the apparatus.

We have chosen to use the data to deduce two values of D and one value of τ. The observed quantities are signal amplitude, S, relative phase angle, θ, and distance from the light spot to the contact, r. Then, following Avery and Gunn (1955), the relevant equations are as follows :

$$D = \frac{\omega_1}{2} \cdot \left(\frac{\partial r}{\partial \theta}\right)_{\omega_1} \cdot \left[\frac{\partial r}{\partial \ln (Sr/r_0)}\right]_{\omega_1} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and

$$\left(\frac{\partial \theta}{\partial r}\right)_{\omega_1}\left[\frac{\partial r}{\partial \ln (Sr/r_0)}\right]_{\omega_2} = \left[\frac{(1+\omega_1{}^2\tau^2)^{1/2}-1}{(1+\omega_2{}^2\tau^2)^{1/2}+1}\right] \quad . \quad . \quad . \quad (3)$$

where r_0 is the radius of the light spot and $\omega_1 = 2\pi$ (4000) and $\omega_2 = 2\pi$ (800) respectively. Independent values of D and τ may be deduced from eqns. (2) and (3). When τ is known a second value of D may be obtained from the following equation :

$$D = \frac{(1+\omega_1{}^2\tau^2)^{1/2}-1}{2\tau(\partial\theta/\partial r)_{\omega_1}{}^2} \, . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In the present experiments a conventional travelling light spot equipment was used, the signal being fed to a phase-sensitive bridge. A phase reference signal for the bridge was derived from a phototransistor in the light beam, which was chopped mechanically. This arrangement ensured that the reference and signal frequencies were coherent regardless of drift in chopper speed.

A schematic diagram of the phase-sensitive bridge and associated circuitry is given in fig. 1. The bridge is balanced by adjustment of the phase of the incoming signal. In a phase-shift circuit of the type shown the phase angle is given by

$$\tan \tfrac{1}{2}\theta = -\omega CR.$$

With a high-μ triode, a maximum phase-shift of about 140° is obtained. For the measurement of signal amplitude the phase is adjusted to give

a maximum reading on the output meter but for the measurement of phase angle considerably greater sensitivity can be obtained by adjusting the signal phase to give zero deflection on the meter.

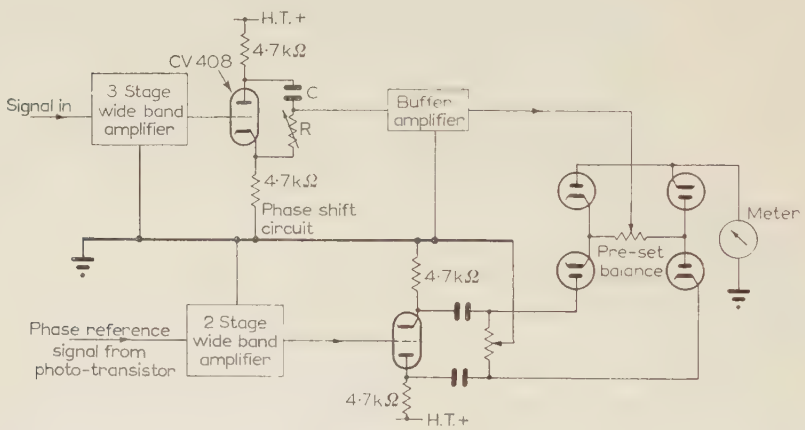
3.2. *The Travelling Light Spot : Results*

Some of the results obtained by the travelling light spot method on a number of N- and P-type germanium crystals are summarized in table 2. Figures 2 and 3 show some phase and amplitude measurements as a function of distance from the detector contact for an N- and P-type crystal respectively. From these results the following generalizations may be made.

(1) In N-type crystals $D_{\perp}$ is equal to the normal value of D but $D_{\parallel}$ is significantly greater. This result is in agreement with the drift mobility data. The apparent carrier lifetime is also generally greater parallel to the dislocations than across them, though this is not always so.

(2) In P-type crystals the values of $D_{\perp}$ and $D_{\parallel}$ are not significantly different from the normal value for electrons and all the anisotropy in diffusion length is due to anisotropy in τ.

Fig. 1



Circuit diagram of phase sensitive bridge.

No significance may be attached to the actual values of $D_{\parallel}$ obtained in the N-type crystals. The crystals used were markedly non-uniform in dislocation density and only regions showing marked anisotropy in diffusion length were selected for detailed examination. Fortunately crystal uniformity is only required over distances ~ 1 mm in this experiment, compared with 7–8 mm in a drift mobility experiment. Short range non-uniformity, indicated by non-linearity of the phase and amplitude curves was observed occasionally, however, particularly in the $\perp$-direction. That these anomalies were a feature of the crystal could be shown by moving the detector contact and remeasuring, when the

anomaly reappeared in the same position of the light spot but at a different radius. As the theory upon which eqns. (2), (3) and (4) are based cannot be applied unless $\partial\theta/\partial r$ and $\partial \ln(Sr/r_0)/\partial r$ are single-valued, measurements of this type were excluded from the results.

Table 2. Values of D and τ deduced from (1) Amplitude at 800 c/s ; (2) Amplitude at 4 kc/s ; (3) Phase at 4 kc/s

N-type crystals

Parallel to dislocations			Across dislocations		
$D_{(2,3)}$	$D_{(1,3)}$	$\tau_{(1,3)}$	$D_{(2,3)}$	$D_{(1,3)}$	$\tau_{(1,3)}$
$cm^2\ sec^{-1}$	$cm^2\ sec^{-1}$	sec	$cm^2\ sec^{-1}$	$cm^2\ sec^{-1}$	sec
116	130	19·3	50	49	10·0
87	82	21	43	41	12·7
89	77	33	46	41	33
109	105	13	—	—	—
110	110	48	78	55	11
81	77	22	without background light		
102	86	14	with background light		
54	66	24	without background light		
83	82	14	with background light		

P-type crystals

$D_{(2,3)}$	$D_{(1,3)}$	τ	$D_{(2,3)}$	$D_{(1,3)}$	τ
88	93	55	99	79	22
89	87	69	100	84	33

§ 4. The Measurement of the Drift Mobility of Holes at High Electric Fields

4.1. *High Field Drift Mobility : Experimental*

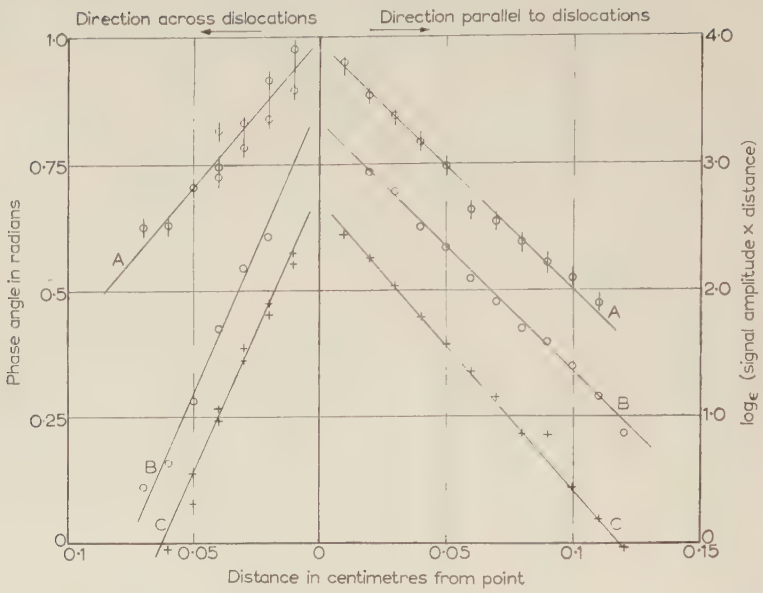
To measure the drift mobility of holes at high fields in N-type filaments cut parallel to and across the dislocations the technique described by Gibson and Granville (1956) was used. Precautions were taken to ensure that :

(1) The injected hole density was kept low enough to reduce conductivity modulation effects to negligible proportions.

(2) Only regions of constant crystal resistivity were selected for examination.

(3) The non-uniformity of carrier lifetime along the filaments, due to variations in dislocation density, did not influence the results significantly.

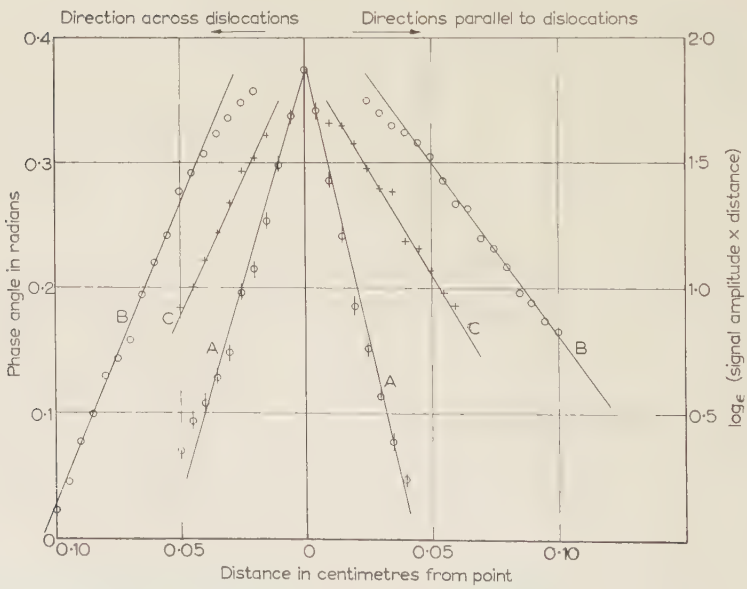
Fig. 2



Signal amplitude and phase as a function of distance from detector contact on N-type germanium.

Curve A. Phase at 4 kc/s.
Curve B. Amplitude at 800 c/s.
Curve C. Amplitude at 4 kc/s.

Fig. 3



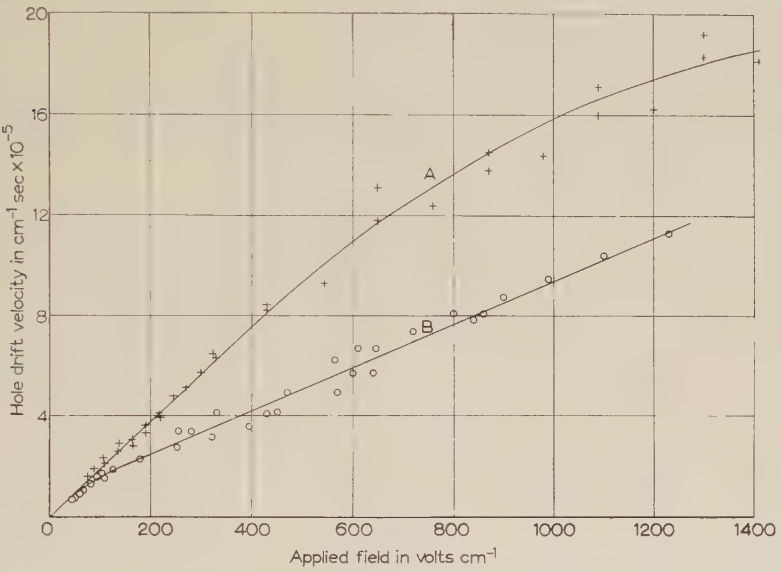
Signal amplitude and phase as a function of distance from detector contact on P-type germanium.

Curve A. Phase at 4 kc/s.
Curve B. Amplitude at 800 c/s.
Curve C. Amplitude at 4 kc/s.

4.2. *High Field Drift Mobility : Results*

It was found that, in N-type germanium the drift mobility of holes parallel to the dislocations was the same as in undislocated material up to 3000 volt cm^{-1}. The drift mobility across the dislocations, on the other hand, was significantly less than the normal value for fields greater than about 100 volt cm [1]. Some typical results are shown in fig. 4. The initial slope corresponds to a drift mobility of 1900 $cm^2 volt^{-1} sec^{-1}$, which is a reasonable value for high resistivity material ($\sim$30 ohm cm) and in agreement with the results obtained by the conventional drift experiment (§ 2).

Fig. 4



Drift velocity of holes in dislocated N-type germanium as a function of applied electric field.

Curve A. Parallel to the dislocation array.
Curve B. Across the dislocation array.

The mobility of the majority carriers can be deduced from current/voltage characteristics of the filaments and no unusual behaviour in either N- or P-type samples was observed up to 4 kv cm^{-1}.

§ 5 Discussion of Results

Shockley (1953) has suggested, on theoretical grounds, that edge dislocations act as a row of acceptor levels due to the dangling bonds. This model has been applied successfully to interpret the Hall effect in dislocated germanium (Pearson *et al.* 1954) and the variation of carrier lifetime with temperature (Wertheim and Pearson 1957). If a significant fraction of the acceptor levels are occupied the potential energy of

electrons is raised locally and the dislocation may be considered as a thread of P-type material embedded in the crystal. The application of the model to the present results will be described quantitatively in a subsequent paper (Gibson and Paige 1958), but a qualitative interpretation of the data will now be given.

5.1. *N-type Germanium*

Threads of P-type material in an N-type bulk crystal will be surrounded by a space-charge region or P–N junction. The junction may be represented by a barrier capacitance and conductance in parallel. The dislocation is therefore analogous to a coaxial line formed by an inner P-type conductor and outer N-type conductor. As a P-type region represents a potential minimum for holes, injected carriers will be captured by the dislocations where they will become majority carriers. If the coaxial line was lossless a space-charge signal would propagate down the line with a velocity near that of light and a hole would be emitted anywhere along the line to compensate for the added charge†. In practice the line is lossy and a hole will be emitted a finite time later at a finite average distance from the point of entry. A captured hole may move in either direction along the dislocation, so that, on the average, its mean position is unaffected by capture in a dislocation 'thread'. It follows therefore that dislocations provide a mechanism for enhanced diffusion of holes along the dislocations without any corresponding increase in hole mobility, so that eqn. (1) is not applicable to this system. Clearly the dislocations cannot assist the diffusion of holes perpendicular to the array.

As already stated, a hole captured in a dislocation thread will, in the absence of an electric field, leave on the average at the point of entry. If an electric field is applied along the dislocation the holes will drift at a velocity determined by the field and their mobility, as in any conductor. It is now convenient to change the frame of reference so that the holes are stationary and the dislocation is moving in the opposite direction. An injected hole, entering the dislocation thread, will now leave, on the average, a distance $\mu.E.t.$ further down the dislocation. Capture in a dislocation thread, therefore, will have no effect on the drift mobility in this direction.

When the electric field is directed across the dislocations a very similar argument may be applied. In the absence of an external field a hole will enter and leave, on the average, at the same point in the cross direction. If the frame of reference is moving in the presence of a field, the hole will behave as if it were drifting at the normal rate. This argument will be valid provided that the distance moved by the dislocation in the time for which the hole is trapped is less than the radius of the thread. This

† This result is essentially the same as that obtained by Moore and Webster (1955) for the propagation of a hole through a 'floating' alloyed P region on N material.

condition cannot be expected to apply at high electric fields and provides the basis for interpreting the data given in § 4.

Apart from the major features of dislocated N-type material discussed above it is possible to fit some of the relatively minor features into the model. For example, it may be expected that the density of holes at the dislocation threads in equilibrium will be increased by background illumination, with a resultant increase in conductivity of the inner P-type conductor and increased effective diffusion constant associated with the dislocation thread (table 2). It is also of interest to note that the high diffusion rate of holes along the dislocations provides an additional mechanism to that suggested by Bell and Hogarth (1957) for enhanced surface recombination in filaments cut across the dislocations.

5.2. *P-type Germanium*

In P-type material the dislocations are, of course, still P-type. However, the negative charge density along the dislocation may be significantly higher than the acceptor density in the high resistivity bulk material so there is still a space-charge region in which the electron potential energy is increased. The presence of this barrier will impede the capture of minority carriers (electrons) by the dislocations and, following the argument given by Bell and Hogarth, provides a mechanism by which τ may be anisotropic. The dislocations will clearly have no significant effect on the diffusion constant of the electrons, however, and no anisotropy of D can be expected.

Acknowledgments

We are indebted to our colleagues for their advice and help and for the supply of suitable crystals. The paper is published by permission of the Controller, H.M. Stationery Office.

References

Avery, D. G., and Gunn, J. B., 1955, *Proc. phys. Soc. Lond.* B, **68,** 918.
Bell, R. L., and Hogarth, C. A., 1957, *J. Electron. and Control.*, **3,** 455.
Gibson, A. F., and Granville, J. W., 1956, *J. Electron.*, **2,** 259.
Gibson, A. F., and Paige, E. G. S., 1958, *Phil. Mag.*, **3,** 950.
Lawrance, R., and Gibson, A. F., 1952, *Proc. phys. Soc. Lond.* B, **65,** 994.
Moore, A. R., and Webster, W. M., 1955, *Proc. Inst. Radio Engrs, N.Y.*, **43,** 427.
Pearson, G. L., Read, W. T., and Morin, F. J., 1954, *Phys. Rev.*, **93,** 666.
Shockley, W., 1953, *Phys. Rev.*, **91,** 228.
Wertheim, G. K., and Pearson, G. L., 1957, *Phys. Rev.*, **107,** 694.

An Interpretation of certain Transport Properties in Germanium containing Parallel Arrays or Edge Dislocations†

By A. F. Gibson and E. G. S. Paige
Royal Radar Establishment, Malvern

[Received April 28, 1958]

Abstract

An interpretation is given of the anisotropic effects observed by Arthur *et al.* (1958) in germanium containing parallel arrays of edge dislocations. The anisotropy of the diffusion constant and high field mobility in N-type crystals is considered quantitatively. The diameter of the space-charge cylinder surrounding the dislocations ($1{\cdot}6 \times 10^{-4}$ cm) and the fraction of time an injected hole spends within the space-charge region ($\frac{1}{2}$) is deduced from the analysis.

§ 1. Introduction

It has been established that edge dislocations in germanium act as acceptor centres (Pearson *et al.* 1954) and that they are efficient sites for recombination (Wertheim and Pearson 1957). Shockley (1953) has accounted for the acceptor type behaviour in terms of the dangling bonds associated with an edge dislocation. Read (1954) has presented a model of an edge dislocation in N-type germanium. He calculates that up to about 10% of the available sites can be occupied by electrons, the electrostatic charge on the dislocation being neutralized by a cylindrical space-charge region. The limitation on the occupancy is imposed by the coulomb interaction energy set up between adjacent electrons on the dislocation.

Observations on the lifetime of minority carriers in dislocated material have not lead to a more detailed model of the dislocation and its effect on the surrounding medium (Wertheim and Pearson 1957). However, recently Bell and Hogarth (1957), using crystals containing parallel arrays of edge dislocations, have found that the diffusion length is anisotropic in both N- and P-type germanium. Further investigations of this effect together with observations on the mobility at high fields have been made by Arthur *et al.* (1958). A summary of the latters experimental observations will be given in § 2. In the remainder of this paper a model will be presented which accounts for the anisotropic effects. In particular, a quantitative analysis is given of the anisotropy of diffusion constant found in dislocated N-type crystals. Certain parameters of the model are calculated from the experimental data of Arthur *et al.* (1958).

† Communicated by the Authors.

§ 2. Summary of Experimental Observations

The model, which we shall consider quantitatively for N-type germanium, has been constructed to explain the observations made by Arthur *et al.* (1958). A summary of the effects observed in crystals containing parallel arrays of edge dislocations follows.

For N-type germanium :

(1) The diffusion constant of holes parallel to the dislocations is greater than that measured perpendicular. The latter has the normal value.

(2) The lifetime of holes is not usually isotropic with direction of diffusion.

(3) The apparent diffusion constant and lifetime of holes parallel to the dislocations increase and decrease respectively in the presence of background illumination.

(4) At low fields the drift mobility of holes is isotropic and has the normal value. At fields in excess of 100 v cm^{-1} the mobility perpendicular to the dislocations becomes significantly less than that parallel to the dislocations—the latter has the usual dependence on field.

(5) The conductivity is isotropic in the same specimens at all fields.

For P-type germanium the diffusion constant of electrons is isotropic but the lifetime is greater for diffusion parallel to the dislocations than across them. The conductivity is again isotropic at all fields.

The magnitude of these effects varied from point to point in both N- and P-type crystals.

§ 3. The Proposed Model

Two major difficulties in analysing the electrical effects of dislocations are that (*a*) the concentration and nature of impurities segregated at dislocations are at present unknown, and (*b*) the density of dislocations are not uniform and they may be polygonized into ' walls '. The variation in magnitude of the anisotropic effects are probably related to the variations in dislocation density. Evidence that impurities segregated at dislocations are not playing a vital role in the anisotropic effects is that anisotropy was observed by Bell and Hogarth (1957) in both plastically deformed and grown dislocated crystals. It is most unlikely that dislocations formed under these two conditions would have the same impurity atmospheres. The model proposed, therefore, is essentially the same as that considered by Read (1954). However, it is found necessary to modify his model in certain respects and in the remainder of this section these modifications are presented.

A qualitative discussion of the anisotropic effects using the proposed model has been given by Arthur *et al.* (1958).

3.1. *The Free Hole Density*

The electrical effects are considered to arise from the trapping of electrons at dangling bonds present at an edge dislocation ; these traps

are hereafter called dislocation sites. The trapped electrons raise the potential energy of electrons in the vicinity of the dislocation and increase the free hole density above that of the bulk. In high resistivity N-type germanium the material becomes P-type in a cylindrical volume around the dislocation, which we call the inversion cylinder. Even in P-type material sufficient electrons can be trapped at dislocation sites to significantly increase the local hole density, forming a P+ cylinder in the P-type crystal.

Read (1954) considered the electrical effects of dislocations in low resistivity N-type germanium. He is able, therefore, to neglect the contribution of free holes to the space-charge surrounding the trapped electrons at sufficiently low temperatures. Since we are concerned with high resistivity (20 Ω cm) material at room temperature we modify Read's model to include a significant free hole contribution to the space-charge region. As will become apparent, the P-type inversion cylinder is vital to the analysis of the anisotropic effects. Direct experimental evidence for the existence of inversion cylinders in one of the crystals used by Arthur *et al.* (1958) has been presented by Hogarth and Baynham (1958). They have observed P-type rectification at ' walls ' formed by the polygonization of dislocations in N-type germanium. They also found a significant lowering of the conductivity near dislocations in P-type material. Tweet (1955) has found relatively high P-type conductivity at the grain boundary in both P- and N-type gold doped bicrystals at low temperatures. The mobility of these holes was not so low as to indicate that at room temperature it would differ appreciably from the normal value.

3.2. *Energy Level of the Dislocation Site*

To provide a mechanism for the anisotropy of lifetime in P-type dislocated crystals Bell and Hogarth (1957) assumed there was a potential barrier to minority carriers surrounding the dislocation. In the proposed model this is formed by the cylindrical PP+ junction. The barrier height will depend on the ratio of the mean separation between ionized acceptors to the average distance between electrons trapped at dislocation sites. As a criterion for an effective barrier we take this ratio to be equal to or greater than unity. Since anisotropic effects have been observed in 12 Ω cm P-type germanium we can use this criterion to estimate an upper limit to ϵ_2, the energy of the dislocation site relative to the top of the valence band. Fermi statistics yield an upper limit of ϵ_2 of 0·4 ev. The use of Fermi statistics will tend to overestimate ϵ_2 but because of the small fraction of dislocation sites occupied the approximation will be fairly good. This is significantly less than the value of 0·5 ev used by Read (1954) in his calculations. The Hall constant and resistivity measurements on plastically deformed 15 Ω cm N-type germanium made by Pearson *et al.* (1954) were consistent with $\epsilon_2 = 0{\cdot}5$ ev. An interpretation of the variation of lifetime with temperature in 2 Ω cm P-type germanium by Wertheim and Pearson (1957) is consistent with the 0·4 ev or smaller

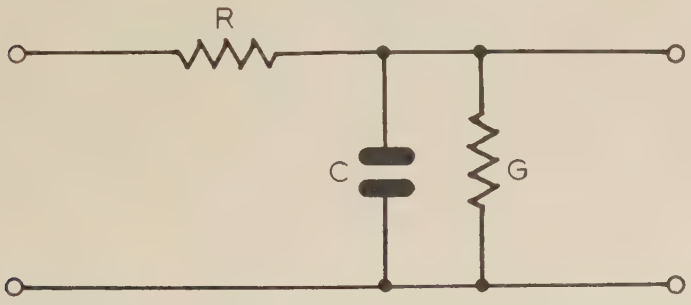
value of ϵ_2. Unfortunately, because the dislocation density is about $10^6\,\mathrm{cm}^{-2}$, no information regarding ϵ_2 would be gained from Hall constant measurements on the crystals used by Bell and Hogarth (1957) and Arthur *et al.* (1958). In the absence of more precise data we take ϵ_2 to be about 0·4 ev in our crystals, and, as previously stated, we do not believe this is associated with a particular impurity atmosphere. It is also significant that it will be found necessary to use this value of ϵ_2 to account for the apparent anisotropy of diffusion constant in N-type germanium.

§ 4. Characteristics of the Inversion Cylinder in N-Type Germanium

In discussing the anisotropy of the diffusion constant in N-type crystals Arthur *et al.* (1958) suggested that the P-type cylinder in N-type material resembled a lossy transmission line. Space-charge signals due to fluctuations in the hole concentration in the inversion cylinder would be propagated along the cylinder with a change of amplitude and phase. In this section we determine these two quantities in terms of characteristics of the line, leaving their relationship to experimental observation till § 5.

The coaxial transmission line will be formed from iterative elements of the form shown in fig. 1. In the figure, R is the resistance per unit length

Fig. 1



Element of transmission line.

of the inner P-type conductor which has a radius r_0. Between the radii r_0 and r_1 is the relatively carrier free region of the PN junction having a barrier capacitance and conductance per unit length of C and G respectively. Outside r_1 the space-charge is zero and the material undisturbed N-type. For such a line the propagation equation of a signal S will be

$$S = S_0 \exp\left(-l.\sqrt{\frac{R}{Z}}\right) \equiv S_0 \exp[-l(\alpha + i\beta)], \quad \ldots \quad (1)$$

where $Z = (G + i\omega C)^{-1}$; ω is the frequency of the fluctuating signal. The

attenuation, α, and the phase shift, β, per unit length, obtained from eqn. (1) are

$$\alpha = \frac{d(\ln S)}{dl} = \left(\frac{RG}{2}\right)^{1/2} \left\{\left(1+\frac{\omega^2 C^2}{G^2}\right)^{1/2}+1\right\}^{1/2}, \quad . \quad . \quad . \quad (2)$$

$$\beta = \frac{d\theta}{dl} = \left(\frac{RG}{2}\right)^{1/2} \left\{\left(1+\frac{\omega^2 C^2}{G^2}\right)^{1/2}-1\right\}^{1/2}, \quad . \quad . \quad . \quad (3)$$

where $\theta = \beta l$ has been introduced for convenience.

It would now appear that we have to calculate, C, G and R. In practice it is only necessary to calculate R and G_p, the hole conductance of the barrier. This is particularly fortunate since R and G_p are determined by the properties of the inner conductor and the undisturbed N-type region respectively. Compared with C, R and G_p are relatively insensitive to the shape and potential distribution in the junction. The barrier conductance per unit length for a small bias will be given by

$$\begin{aligned} G &= G_p + G_n \\ &= (I_{p_0} + I_{n_0})q/kT. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4) \end{aligned}$$

Here I_{p_0} and I_{n_0} are the saturation currents per unit length of the junction for holes and electrons respectively. It will become apparent in § 5 that we are only interested in the hole current. The calculation of I_{p_0} has been performed in the Appendix : from the result we obtain a hole conductance per unit length of

$$G_p = 2\pi q \mu_p p_n [\ln(2L/r_1) - \gamma]^{-1}. \quad . \quad . \quad . \quad . \quad . \quad (5)$$

μ_p is the normal hole mobility, p_n the equilibrium density of holes in the N-type crystal, and L is the diffusion length of holes.

The resistance per unit length of the inversion cylinder may be estimated as follows. If a fraction f of dislocation sites are occupied and the separation between sites is a, then a negative change of fq/a resides along unit length of the edge dislocation. For charge neutrality,

$$-\frac{fq}{a} + Pq + \pi r_1^2 (N_D - N_A) q = 0 \quad . \quad . \quad . \quad . \quad (6)$$

where P is the number of holes per unit length of the dislocation and $(N_D - N_A)$ is the density of uncompensated donors. For a hole mobility of μ_{pd} along the cylinder, the resistance per unit length becomes

$$R = \left\{\left[\frac{f}{a} - \pi r_1^2 (N_D - N_A)\right] q \mu_{pd}\right\}^{-1}. \quad . \quad . \quad . \quad . \quad (7)$$

§ 5. The Relationship between the Experimental Observations and the Parameters of the Inversion Cylinder

The travelling light spot and drift mobility experiments of Arthur *et al.* (1958) will be compared with parameters of the inversion cylinder. Then, using the experimental data from the mobility experiments, a value of the apparent diffusion constant in the inversion cylinder together with

the fraction of time spent by an injected hole in the cylinder will be found. These will then be compared with the observed diffusion constants.

5.1. *The Travelling Light Spot Experiment*

The technique used by Arthur *et al.* (1958) which enabled a simultaneous measurement of diffusion constant, D, and lifetime, τ, to be made has been described by Avery and Gunn (1955). The equations from which D and τ can be deduced are

$$\frac{d[\ln(rS/r_0)]}{dr} = (2D\tau)^{-1/2}\{(1+\omega^2\tau^2)^{1/2}+1\}^{1/2}, \quad . \quad . \quad . \quad (8)$$

and

$$\frac{d\theta}{dr} = (2D\tau)^{-1/2}\{(1+\omega^2\tau^2)^{1/2}-1\}^{1/2}, \quad . \quad . \quad . \quad (9)$$

where ω is the modulation frequency of the light spot. These two equations are identical in form to eqns. (2) and (3). The replacement of $d/dl(\ln S)$ by $d/dr(\ln rS/r_0)$ occurs because eqns. (2) and (3) are essentially one dimensional. In practice the variation of S with r is so rapid that $d(\ln S)/dl$ and $d/dr[\ln(Sr/r_0)]$ are very nearly equal. We may therefore define

$$D_d = (RC)^{-1} \text{ and } t_d = CG^{-1}$$

as the effective diffusion constant and release time of a hole in an inversion cylinder. The quantity t_d should not be confused with a carrier lifetime; it is the average time a hole spends within a cylinder. The length $\sqrt{(D_d t_d)}$ represents the mean range of a hole along the cylinder.

Two processes limit t_d, recombination at the dislocation sites and escape to the undisturbed N-type material. Thus the probability per unit time that an excess hole is removed from the inversion cylinder is

$$\frac{1}{t_d} = \frac{1}{t_r} + \frac{1}{t_\epsilon}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where t_r and t_ϵ are average times spent in the inversion cylinder before recombination and escape respectively. Suppose a group of holes have entered the inversion cylinder. Due to their presence the PN junction is forward biased and some holes will be ejected from and some electrons will flow into the P-region. The former process will be determined by the hole conductance of the barrier, the latter by the electron conductance. Accordingly we write

$$t_\epsilon = C/G_p \text{ and } t_r = C/G_n. \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

The fraction of time spent by a hole in an inversion cylinder, η, is simply $t_d(t_c+t_d)^{-1}$ where t_c, the mean time between escape from and re-entering the cylinder, is given by

$$t_c = (\rho_d \cdot 2r_1 \cdot \sqrt{\tfrac{2}{3}} \cdot v_T)^{-1}. \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Here ρ_d is the density of dislocations and v_T is the thermal velocity of the hole.

We may define a minority carrier lifetime in terms of t_r by the equation

$$\tau = t_r \cdot \eta^{-1}. \qquad (13)$$

5.2. *The Drift Mobility Experiment*

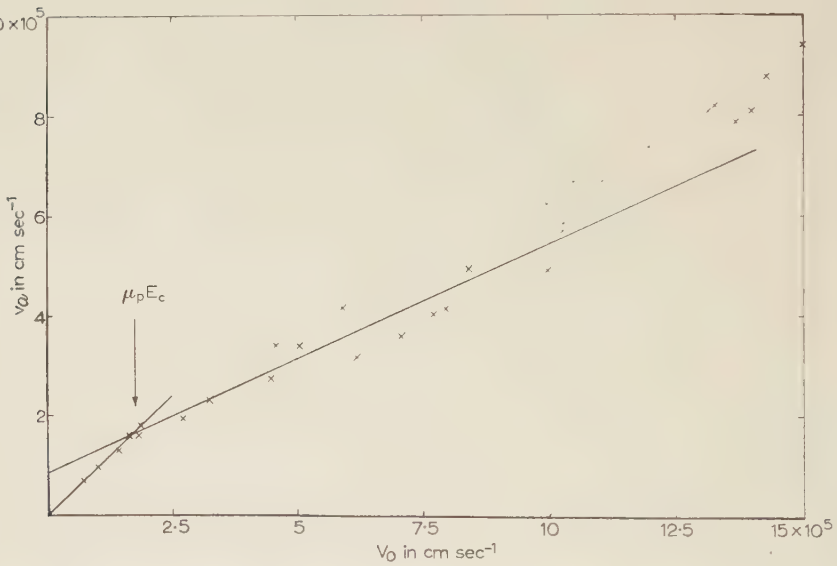
A characteristic time, t_ϵ, has been associated with the average interval between entering and leaving an inversion cylinder. We may therefore look upon the inversion cylinder as a trap of large physical dimensions. If the electric field is perpendicular to the dislocation array and the drift velocity, v, such that $v < \frac{1}{2}(2r_1/t_\epsilon)$, then trapping effects will not be noticeable. However, for fields in excess of E_c, given by

$$E_c = \frac{1}{2\mu_p} \cdot \left(\frac{2r_0}{t_\epsilon}\right), \qquad (14)$$

trapping will lower the apparent mobility. Under these conditions the apparent drift velocity v_a will be

$$v_a = (r_1/t_\epsilon)\eta' + v_0(1-\eta'), \qquad (15)$$

Fig. 2



Variation of the drift velocity perpendicular to the dislocation array (v_a) with the normal drift velocity (v_0) in the same electric field.

where v_0 is the normal drift velocity at the field considered, and $\eta' = t_\epsilon(t_c + t_\epsilon)^{-1}$. At sufficiently high fields, such that $Er_1 \simeq kT/q$, both r_1 and t_ϵ will become field dependent in such a way as to tend to increase v_a to the normal value.

In fig. 2 v_a has been plotted against v_0 and the three drift velocity regions can be seen. As predicted at the lowest fields, there is a linear

region of slope unity ; at higher fields the plot is linear and of slope $(1-\eta')$; at the highest fields, the departure from this slope is observed due to changes in r_1 and t_ϵ.

For an electric field parallel to the dislocation array, each dislocation considered to be of infinite length, the apparent mobility, μ_a, is

$$\mu_a = \mu_{pd}\eta' + \mu_p(1-\eta'). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

5.3. *Parameters Derived from the Experimental Results*

By comparing fig. 2 with eqns. (14) and (15) we obtain $1-\eta' = 0{\cdot}45$, $r_1/t_\epsilon = 1{\cdot}7 \times 10^5\,\mathrm{cm\,sec^{-1}}$ and $(r_1/t_\epsilon)\eta' = 0{\cdot}9 \times 10^5\,\mathrm{cm\,sec^{-1}}$. The two values of η' are consistent and approximately 0·5 thus, from the definition of η', t_c and t_ϵ must be about equal. From etch pit counts ρ_d was found to be approximately $1 \times 10^6\,\mathrm{cm^{-2}}$. Inserting this value in eqn. (12), $r_1 t_c\,(=r_1 t_\epsilon)$ can be calculated. From $r_1 t_\epsilon$ and $r_1 t_\epsilon^{-1}$ we obtain $r_1 = 0{\cdot}8 \times 10^{-4}\,\mathrm{cm}$ and $t_\epsilon = t_c = 5 \times 10^{-10}\,\mathrm{sec}$. The mobility measured parallel to the dislocations has the same value as in relatively dislocation free material at all fields. Because η' is a sufficiently large fraction this result enables us to deduce from eqn. (16) that $\mu_{pd} \simeq \mu_d$.

A value of D_d will be deduced now using the parameters that have been determined for the drift mobility experiments. The resistivity of the crystal used is about $20\,\Omega\,\mathrm{cm}$, hence $(N_D - N_A)$ is $1 \times 10^{14}\,\mathrm{cm^{-3}}$, p_n is $3 \times 10^{12}\,\mathrm{cm^{-3}}$ and the position of the fermi level is 0·42 ev above the valence band. The hole conductance of the barrier is evaluated by substituting the typical value of L, $2 \times 10^{-2}\,\mathrm{cm}$, in eqn. (5). G_p is not sensitive to values of r_1 and L provided $L \gg r_1$, as is the case. Using the value of ϵ_2 suggested in § 3, f is estimated to be close to the limiting value of 0·1. We can now calculate R from eqn. (7) remembering the separation between dislocation sites is 4 Å for an edge dislocation and that equating μ_{pd} to μ_d has been justified. We find $R = 1{\cdot}2 \times 10^{10}\,\Omega\,\mathrm{cm^{-1}}$. Hence a value of the effective diffusion constant in the inversion cylinders is obtained

$$D_d = \frac{1}{RC} = \frac{1}{RG_p t_\epsilon} = 200\,\mathrm{cm^2\,sec^{-1}}.$$

From eqn. (13), and the observed lifetime, it is readily shown that $t_r \gg t_\epsilon$. Also it has been found that injected holes spend about half their life in inversion cylinders. Without a rigorous calculation it is clear that the observed value of the diffusion constant parallel to the dislocations will be intermediate between D_d and the normal value. Perpendicular to the dislocations the inversion cylinders will not affect the diffusion process and the normal value of D, $47\,\mathrm{cm^2\,sec^{-1}}$, is to be expected. Typical results presented by Arthur *et al.* (1958) are :

D parallel to dislocations $\simeq 100\,\mathrm{cm^2\,sec}$,

D perpendicular to dislocations $\simeq 45\,\mathrm{cm^2\,sec}$.

If a value of ϵ_2 of 0·5 ev is taken instead of 0·4 ev, f falls appreciably and $D_{//} \to 0$. It is important to notice that the observed value of D will be sensitive to the dislocation density since this determines η'. This is the basis for the statement in § 3, that the variation of the anisotropic effects from point to point in the crystal was due to ρ_d.

§ 6. Conclusion

Arthur *et al.* (1958) have shown that an anisotropy of the diffusion length was due to an anisotropy in τ in P-type crystals while an anisotropy in D was always found in N-type crystals. The model proposed in § 3 has been used to give a qualitative explanation of the former and a quantitative explanation of the latter. It does not provide a mechanism for an anisotropic D in P-type crystals, since the transmission line analogy is no longer valid, and indeed no anisotropy of D is observed.

An explanation of the drift mobility data of holes in N-type specimens has led to a calculation of the diameter of the space-charge cylinder surrounding the dislocation. The value of $1{\cdot}6 \times 10^{-4}$ cm for the diameter is not inconsistent with the observations of Hogarth and Baynham (1958) since they probed across dislocations which had been polygonized into a 'wall'. From R, r_1 and $\rho_{//}$ the conduction of the inner P-type conductors parallel to the dislocations can be calculated. It is found that no significant enhancement of the conductivity parallel to the dislocation should be observed. This is consistent with the experimental measurements on the dependence of the conductivity on the direction of the electric field.

The model for N-type germanium that has been presented does not lead to an anisotropy in τ which sometimes accompanies the anisotropy in D. Since the measured lifetime is typically 10 μsec, an injected hole passes in and out of an inversion cylinder 10^4 times before it recombines with an electron. Clearly, then, the probability of recombination will be independent of the direction of diffusion relative to the dislocation array. The increased D measured parallel to the dislocation in the presence of a background light arises from a decrease in R due to the free carrier concentration in the inversion cylinder rising. The lifetime is limited by the rate at which electrons can surmount the potential barrier surrounding the dislocation sites. The photovoltage set up by the background light will assist the process and decrease the lifetime.

So far the possibility of polygonization of the dislocations into dislocation 'walls' has not been considered. Extension of the calculation of § 4 and § 5 to a polygonized crystal has been performed and an enhanced value of D parallel to the walls can be obtained. However, the introduction of new variable parameters (e.g. packing of dislocations in the wall, dimensions of the wall) in this analysis makes it less stringent than that which we have given.

APPENDIX

We wish to calculate the hole conductance of a cylindrical PN junction If a small forward bias is applied to the junction, holes are injected into the N-type material and recombine at the randomly distributed recombination sites and at the dislocations parallel to the cylinder. Treating the problem as though the recombination sites associated with the dislocation were also randomly distributed we can write the continuity equation in cylindrical coordinates as

$$\frac{d^2p(r)}{dr^2}+\frac{1}{r}\frac{dp(r)}{dr}-\frac{p(r)}{L^2}=0. \quad . \quad . \quad . \quad . \quad . \quad (A.1)$$

Here L is the diffusion length of holes and $p(r)$ the hole density in the N region. Substituting $x=r/L$ and introducing $v(x)\equiv p(r)$, eqn. (A.1) becomes

$$\frac{x^2d^2v(x)}{dx^2}+\frac{x\,dv(x)}{dx}-x^2v(x)=0. \quad . \quad . \quad . \quad . \quad (A.2)$$

The solution of this equation (Jeffreys and Jeffreys 1956) is

$$v(x)=\alpha I_0(x)+\beta K_0(x),$$

where

$$I_0(x)=\sum_{k=0}^{\infty}(\tfrac{1}{2}x)^{2k}.\,(k\,!)^{-2},$$

$$K_0(x)=-I_0(x)\ln(\tfrac{1}{2}x)+\sum_{k=0}^{\infty}(\tfrac{1}{2}x)^{2k}(k\,!)^{-2}F(k)$$

and $F(k)$ is the digamma function.

Using the boundary conditions that $v=0$ at $x=\infty$ and $v=v_1$ at $x=x_1=r_1/L$ and making approximations valid because $r_1/L\ll 1$, the constants α and β are found,

$$\alpha=0 \text{ and } \beta=v_1[\ln(2L/r_1)-\gamma]^{-1},$$

where $\gamma=0{\cdot}58$.

Therefore, for values of r such that $r/L\ll 1$,

$$p(r)=-p_1\left[\ln\frac{2L}{r_1}-\gamma\right]^{-1}.\left(\ln\frac{r}{2L}+\gamma\right). \quad . \quad . \quad . \quad (A.3)$$

The diffusion current per unit length of the cylindrical junction is

$$i_p=-q\,.\,2\pi r\,.\,D_p\frac{dp(r)}{dr}$$

$$=q\,.\,2\pi D_p\,.\,p_1[\ln(2L/r_1)-\gamma]^{-1} \quad . \quad . \quad . \quad . \quad (A.4)$$

where D_p is the diffusion constant of holes. If a forward bias of V volts is applied such that $V\ll kT/q$, then following the usual PN junction theory, we can write the conductance per unit length due to holes as

$$G_p=2\pi q\mu_p p_n\{\ln(2L/r_1)-\gamma\}^{-1}. \quad . \quad . \quad . \quad . \quad (A.5)$$

Here p_n is the thermal equilibrium concentration of free holes in the N region and μ_p the hole mobility.

In this approach we have treated the recombination sites associated with all dislocations as if they were randomly distributed. Consider the parallel dislocaions to be arranged on a square network. then the eight

nearest neighbour parallel dislocations, at about a radius, R, of $(2\rho_d)^{-1/2}$, may dictate the conductance by forming a 'recombination surface' around the dislocation. Under these conditions eqn. (A.1) is modified by replacing L by L_0 (the diffusion constant in the absence of the parallel dislocation array) and by changing the boundary condition $v=0$ at $x=\infty$ to

$$sv(x) = -\frac{D}{L_0} \cdot \frac{dv(x)}{dx} \quad \text{at} \quad x = \frac{R}{L_0},$$

s is the recombination velocity of a surface containing eight parallel dislocations in $2\pi R$; its value can be estimated from Okada's (1955) data.

The expression for the hole conductance per unit length is

$$G_p' = 2\pi q \mu_p p_n \frac{Rs}{D}.$$

Since G_p' is nearly two orders of magnitude less than G_p calculated from eqn. (A.5) the neighbouring parallel dislocations do not have a controlling influence on the conductance, a result we expect since otherwise $L \sim R$. We are justified, therefore, in using (A.5) except that this is a d.c. conductivity, an a.c. conductivity is appropriate for our problem but for small bias they are approximately equal.

Acknowledgments

We wish to thank our colleagues for helpful discussions, particularly Mr. A. C. Prior. The paper is published by permission of the Controller, H.M. Stationery Office.

References

Arthur, J. B., Gibson, A. F., Granville, J. W., and Paige, E. G. S., 1958, *Phil. Mag.* **3**, 940.
Avery, D. G., and Gunn, J. B., 1955, *Proc. phys. Soc. Lond.* B, **68**, 918.
Bell, R. L., and Hogarth, C. A., 1957, *J. Electron. and Control*, **3**, 455.
Hogarth, C. A., and Baynham, A. C., 1958, *Proc. phys. Soc. Lond.*, **71**, 647.
Jeffreys, H., and Jeffreys, B., 1956, *Method of Mathematical Physics*. 3rd *ed.* (Cambridge: University Press).
Okada, J., 1955, *J. phys. Soc. Japan*, **10**, 1110.
Pearson, G. L., Read, W. T., and Morin, F. J., 1954, *Phys. Rev.*, **93**, 666.
Read, W. T., 1954, *Phil. Mag.*, **45**, 775.
Shockley, W., 1953, *Phys. Rev.*, **91**, 228.
Tweet, A. G., 1955, *Phys. Rev.*, **99**, 1182.
Wertheim, G. K., and Pearson, G. L., 1957, *Phys. Rev.*, **107**, 694.

Cosmic Rays in the Earth's Magnetic Field†

By P. Rothwell
Physics Department, Imperial College, London

[Received May 16, 1958]

Abstract

It is shown that the values of cosmic ray cut-off momenta in the earth's magnetic field, observed at many different places, are generally close to the values calculated from Störmer's theory for the motion of charged particles in a dipole field, if the usual centre dipole of the earth is replaced in the Störmer equation by a dipole whose magnitude and direction are determined by the surface field at the place considered. An empirical expression for the actual cut-off momenta (in terms of the centre-dipole field, and the ' surface field ' cut-off momenta) is deduced from the variation in sea-level nucleon intensity between London and Cape Town, and gives good agreement with experimental results over a wide range of latitidues and longitudes. It is concluded that discrepancies between centre dipole predictions and experimental observations of cosmic ray intensities and cut-off momenta are due to differences between the earth's real field and the dipole approximation to it, rather than to distortion of the earth's outer magnetic field by ionized interplanetary matter.

The cut-off momenta and the intensities of cosmic rays measured at various latitudes and longitudes are often not in good agreement with the prediction of Störmer theory for the motion of charged particles in the earth's dipole field. These discrepancies have been ascribed to distortion of the earth's outer magnetic field by highly ionized interplanetary matter (Simpson *et al.* 1956). On the other hand, recent measurements (Rothwell and Quenby 1957) have shown that cosmic ray intensity and surface magnetic field anomalies occur near the same places. This result implies that 'anomalies' in the earth's real field modify the cut-off momenta of cosmic ray particles (of which the intensity is a function). A calculation of the effect on the Störmer cut-off momentum of a given regional magnetic anomaly, would be of great mathematical complexity, so we have used a phenomenological approach to the problem.

Störmer's theory for the motion of charged particles in a dipole field gives the minimum momentum p_D, required for an incoming vertical particle of charge Ze, to reach the earth, as

$$p_D = \frac{Ze}{4c}\frac{M}{R^2}\cos^4\lambda \qquad \ldots\ldots\ldots \quad (1)$$

where λ is the geomagnetic latitude, M is the moment of the earth's dipole, and R is the radius of the earth ($6{\cdot}4 \times 10^8$ cm).

---

† Communicated by the Author.

This may be conveniently expressed in terms of the horizontal field, $H_D = MR^{-3}\cos\lambda$ and the magnetic dip δ_D, given by $\tan\delta_D = 2\tan\lambda$, as

$$p_D = \frac{Ze}{4c} R \frac{H_D}{(1+\frac{1}{4}\tan^2\delta_D)^{3/2}}. \quad \ldots\ldots\ldots \quad (2)$$

We define p_S as the momentum cut-off corresponding to the surface field values of the horizontal component and dip, H_S and δ_S in the eqn. (2) that is

$$p_S = \frac{Ze}{4c} R \frac{H_S}{(1+\frac{1}{4}\tan^2\delta_S)^{3/2}}. \quad \ldots\ldots\ldots \quad (3)$$

In this expression, the usual centred dipole, as determined from measurements of the earth's field over the whole of the earth's surface, has been replaced, effectively, by a dipole whose magnitude and direction are determined by the surface field at the place considered only. The path of a cosmic ray particle in earth's field is, of course, determined not only by the field at the earth's surface, but also by the field further out, which in general approaches the dipole field approximation in value with increasing distance from the earth. The actual vertical cut-off momentum p_C may therefore be expected to lie somewhere between p_S and p_D.

The cut-off momenta of α-particles have been measured directly at a number of places over North America, Europe and Australia. Table 1 compares the observed values of cut-off rigidity p/Z with values computed (*a*) for the dipole field, p_D/Z, and (*b*) for the surface field, p_S/Z. It can be seen that while discrepancies between the observed and the centre dipole values are large in some places, remarkably good agreement is found nearly everywhere between the observed and the 'surface field' values of cut-off momentum.

Direct measurements of cut-off momenta have not been made at latitudes lower than $\sim 40°$; in this region the experimental data obtained with a neutron intensity monitor on a sea voyage between London and Cape Town (Rothwell and Quenby 1957) has been used to find an empirical expression for the cut-off momentum p_C in terms of p_S and p_D: for similar intensities should be observed at places with the same cut-off momentum.

Figure 1 shows the variation of cosmic ray intensity between London and Cape Town (*a*) with p_D and (*b*) with p_S. In (*a*) intensities in the southern hemisphere are higher than in the northern hemisphere, while in (*b*) they are lower.

The condition that similar intensities should be observed at places with the same true cut-off is satisfied if we write

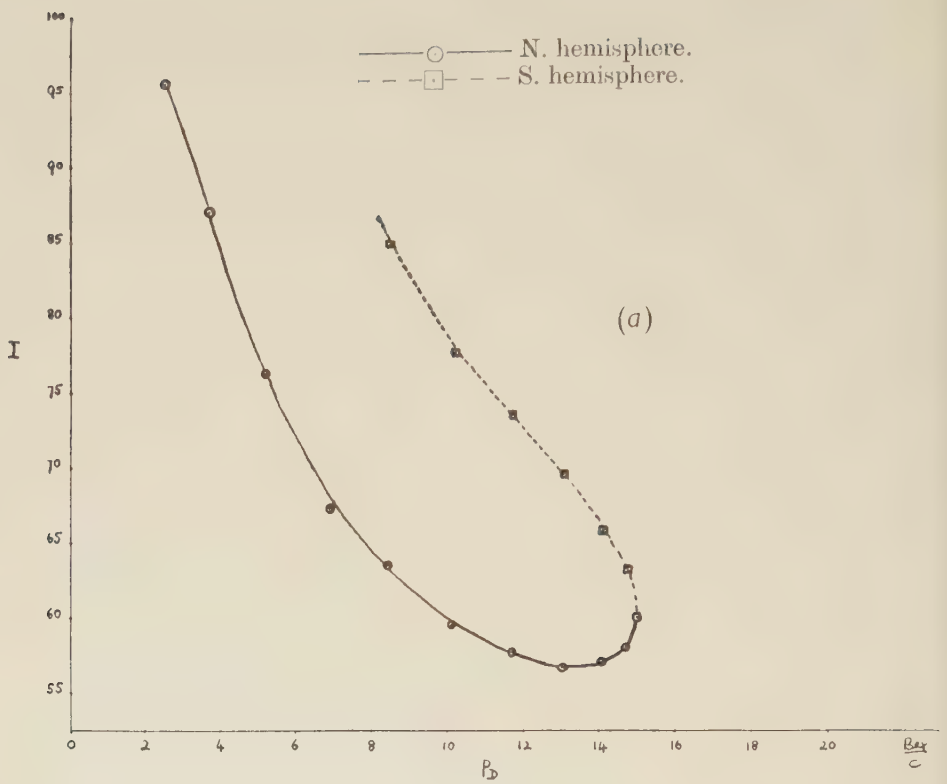
$$p_C = x\,p_S + (1-x)p_D\,(p_D/p_S) \quad \ldots\ldots\ldots \quad (4)$$

where x is a factor which, *if assumed constant over the range of measurements*, can be determined by equating values of p_C at places with similar observed cosmic ray intensities. Table 2 gives values of p_D, p_S and p_D/p_S, for seven pairs of points at which similar intensities were recorded on the voyage of the 'Roxburgh Castle' between Cape Town and London, and the corresponding value of x calculated for each pair. It can be

Table 1. Comparison of Observed and Calculated α-Particle Cut-Off Rigidities

Place of observation and observers	λ (geomag.)	Observed $\frac{p}{Z}$	% Error of observation	(a) $\frac{p_D}{Z}$	$\frac{p_D - p}{p}$	(b) $\frac{p_S}{Z}$	$\frac{p_S - p}{p}$	(c) $\frac{p_C}{Z}$	$\frac{p_C - p}{p}$
Minneapolis (Fowler *et al.* 1957)	55·1° N	$<1{\cdot}1$	$\pm 14\%$	1·6	$>+50\%$	$0{\cdot}8_5$	$<-25\%$	$1{\cdot}1_5$	$>+5\%$
Waukon (Webber private comm.)	53·5° N	1·2	$\pm 12\%$	1·9	$+58\%$	1·1	-8%	$1{\cdot}3_5$	$+12\%$
Grand Rapids (MacDonald 1957)	53·2° N	1·1	$\pm 14\%$	$1{\cdot}9_5$	$+77\%$	$1{\cdot}0_5$	-5%	1·4	$+27\%$
Iowa (MacDonald 1957)	52·4° N	$1{\cdot}3_5$	$\pm 11\%$	$2{\cdot}0_5$	$+48\%$	$1{\cdot}3_5$	0%	$1{\cdot}5_5$	$+15\%$
Missouri (Webber private comm.)	50·6° N	1·6	$\pm 12\%$	$2{\cdot}4_5$	$+53\%$	$1{\cdot}8_5$	$+16\%$	$1{\cdot}9_5$	$+22\%$
San Angelo (Webber private comm.)	41·4° N	4·7	$\pm 6\%$	$4{\cdot}6_5$	-1%	$5{\cdot}0_5$	$+7\%$	$4{\cdot}9_5$	$+5\%$
Bristol (Waddington 1956)	54·6° N	$2{\cdot}5_5$	$\pm 10\%$	1·7	-33%	$2{\cdot}4_5$	-4%	2·3	-10%
N. Italy (Fowler *et al.* 1956)	46·0° N	4·6	$\pm 5\%$	$3{\cdot}6_5$	-21%	4·3	-6%	$4{\cdot}1_5$	-10%
Cagliari (Aly *et al.* 1957) (De Marco *et al.* 1956)	40·3° N	5·6 6·6	$\pm 4\%$ $\pm 6\%$	$5{\cdot}1_5$	-10% -25%	6·3	$+12\%$ -4%	$6{\cdot}0_5$	$+8\%$ -8%
Melbourne (Hooper *et al.* 1958)	47·0° S	2·7	$\pm 15\%$	$3{\cdot}2_5$	$+20\%$	2·7	0%	2·8	$+4\%$

Fig. 1


N. hemisphere.
S. hemisphere.
(a)
I
100
95
90
85
80
75
70
65
60
55
0
2
4
6
8
10
12
14
16
18
20
Bay
c
P_D

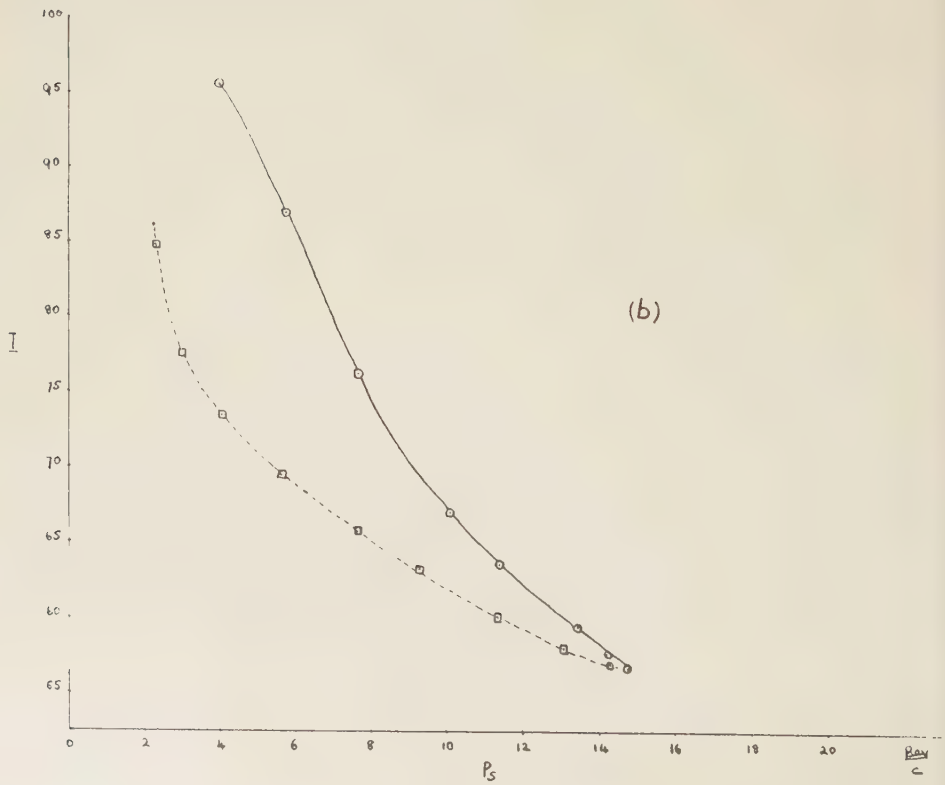
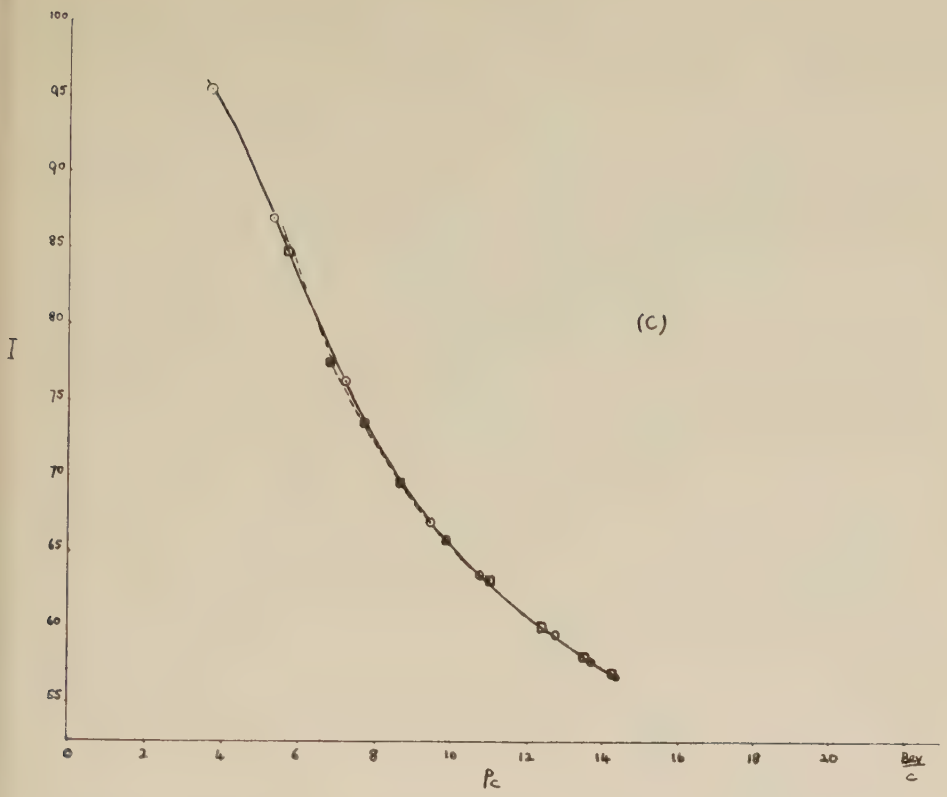

(b)
I
100
95
90
85
80
75
70
65
60
65
0
2
4
6
8
10
12
14
16
18
20
Bay
c
P_S

Fig. 1 (*continued*)



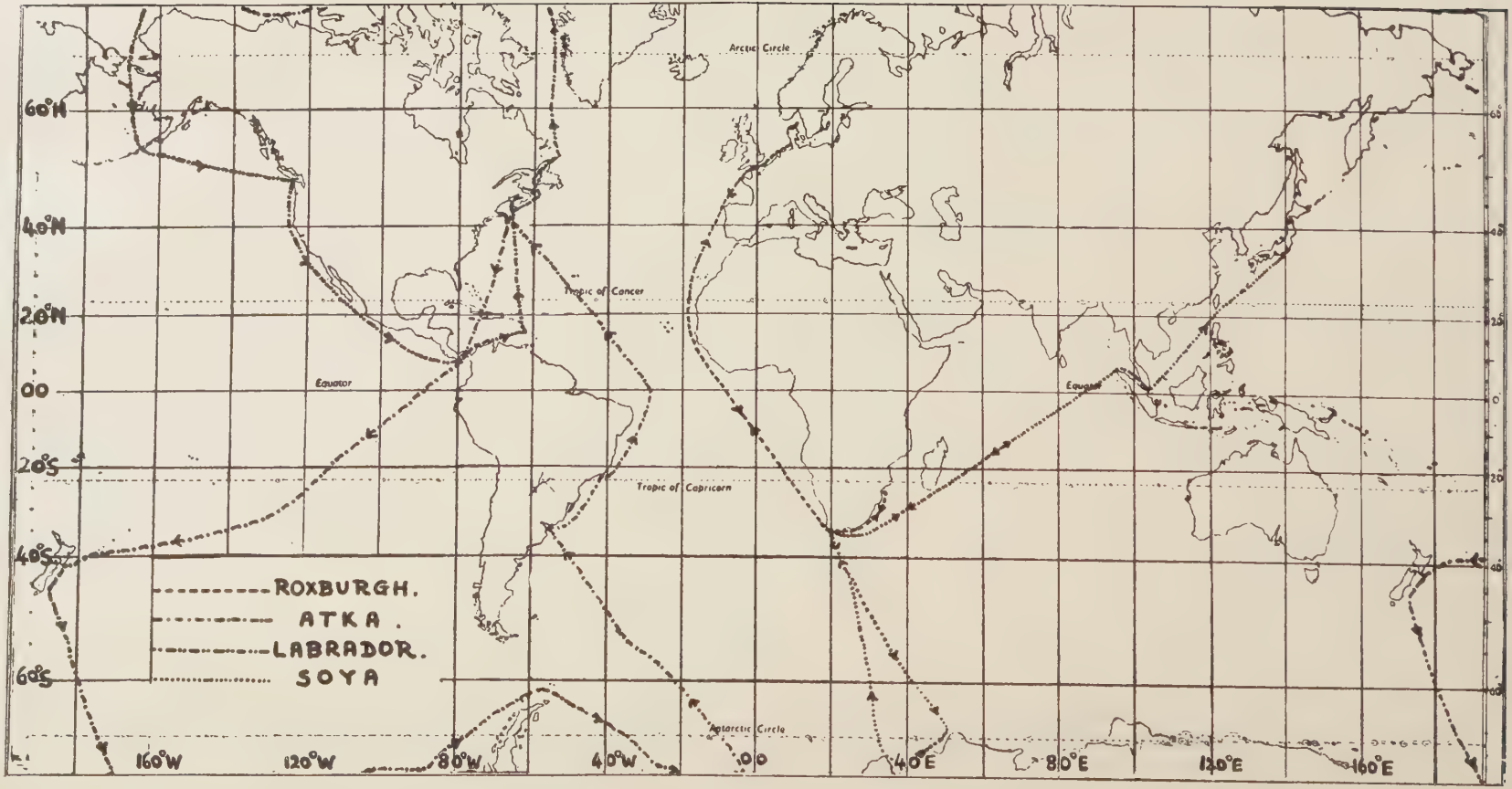
Cosmic ray intensity versus cut-off rigidity on voyage of ' Roxburgh Castle '.

Table 2. Comparison of calculated cut-off momenta p_D, p_S, for cosmic ray particles, at places with similar observed intensities between Cape Town and London, and deduction of factor 'x' in the empirical expression for the actual cut-off momenta

$$p_C = x p_S + (1-x) p_D (p_D/p_S)$$

Intensity	λ	H_S	δ_S	p_D	p_S	$\frac{p_D}{p_S}$	x
85	45° N	$0{\cdot}24_2$	56°	3·9	$6{\cdot}0_5$	$0{\cdot}6_4$	$0{\cdot}88_5$
85	30° S	$0{\cdot}13_5$	$-63{\cdot}0_5$	8·4	2·3	$3{\cdot}6_5$	
80	42° N	$0{\cdot}25_2$	53°	$4{\cdot}5_5$	7·0	$0{\cdot}6_5$	0·88
80	27° S	$0{\cdot}14_5$	−62°	9·6	2·7	$3{\cdot}5_3$	
75	39° N	$0{\cdot}26_5$	50°	5·4	7·9	$0{\cdot}6_8$	0·88
75	22° S	$0{\cdot}16_0$	−59°	11·1	3·5	$3{\cdot}1_8$	
70	36° N	$0{\cdot}27_7$	46°	6·4	9·3	$0{\cdot}7_0$	$0{\cdot}87_5$
70	16° S	$0{\cdot}19_0$	−53°	12·8	5·3	$2{\cdot}4_0$	
65	32° N	$0{\cdot}29_0$	40°	7·8	10·8	$0{\cdot}7_2$	0·88
65	9° S	$0{\cdot}22_3$	−43°	14·3	8·0	$1{\cdot}7_9$	
60	26° N	$0{\cdot}30_5$	30°	9·9	13·1	$0{\cdot}7_6$	$0{\cdot}87_5$
60	0°	$0{\cdot}26_0$	−27°	15·0	$11{\cdot}3_5$	$1{\cdot}3_2$	
57·5	20° N	$0{\cdot}31_2$	18°	11·8	14·4	$0{\cdot}8_2$	0·88
57·5	7° N	$0{\cdot}29_0$	−15°	14·5	13·6	$1{\cdot}0_7$	

Fig. 2



Routes of ships which have recorded cosmic ray neutron intensities.

seen that when p_C is expressed in the form given by eqn. (4), the value of x is, in fact, approximately constant for all latitudes given in the table. Taking $x \simeq 0{\cdot}9$, eqn. (4) may be rewritten as

$$p_C = 0{\cdot}9 p_S + 0{\cdot}1 (p_D^2/p_S). \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Figure 1 (*c*) shows the variation of cosmic ray intensities with p_C; it can be seen that in this case the intensities measured in the northern and southern hemispheres lie on the same curve.

The validity of eqn. (5) may be checked at many different latitudes and longitudes.

(1) Neutron intensities have been recorded on voyages around the east and west coasts of America, and across the Pacific and Indian Oceans (Simpson *et al.* 1956, Rothwell and Quenby 1957, Kodama and Miyazaki 1957) (see fig. 2). Variation of intensity with cut-off momentum calculated (*a*) for the dipole field, (*b*) for the surface field and (*c*) from eqn. (5) are shown in fig. 3. It can be seen, while there are wide variations of intensity with centre dipole cut-off momentum p_D, that similar intensities are observed at places with the same cut-off momentum p_C, over a wide range of latitudes and longitudes, except perhaps in the region of anomalously high horizontal magnetic intensity, crossed by the Soya, south of Japan (Kodama and Miyazaki 1957).

Fig. 3

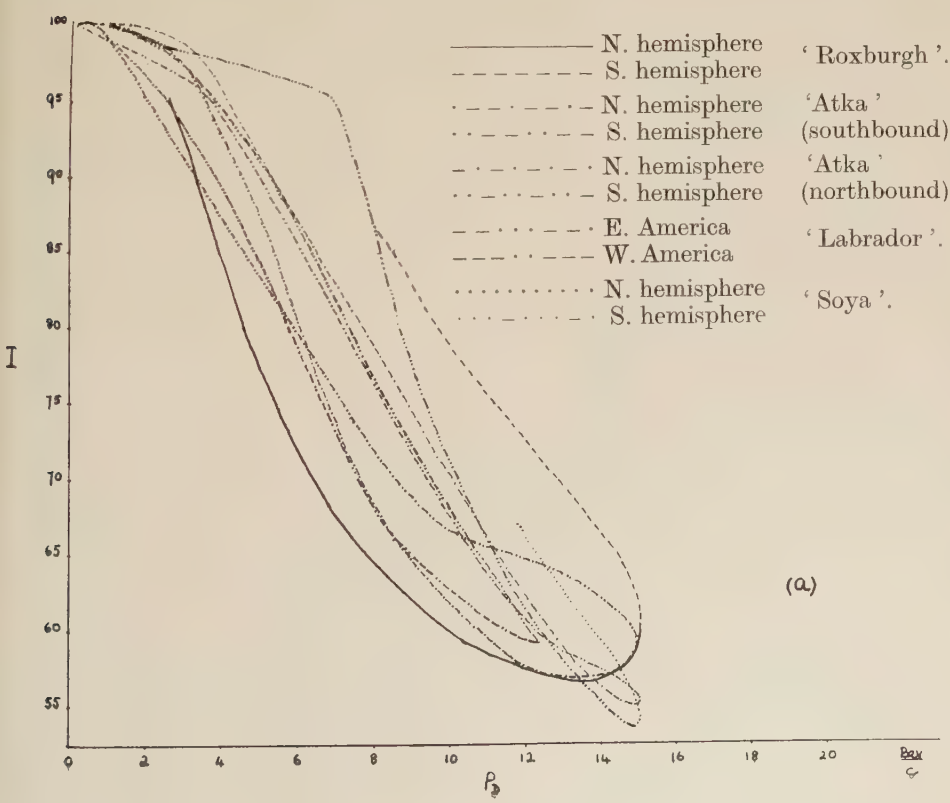
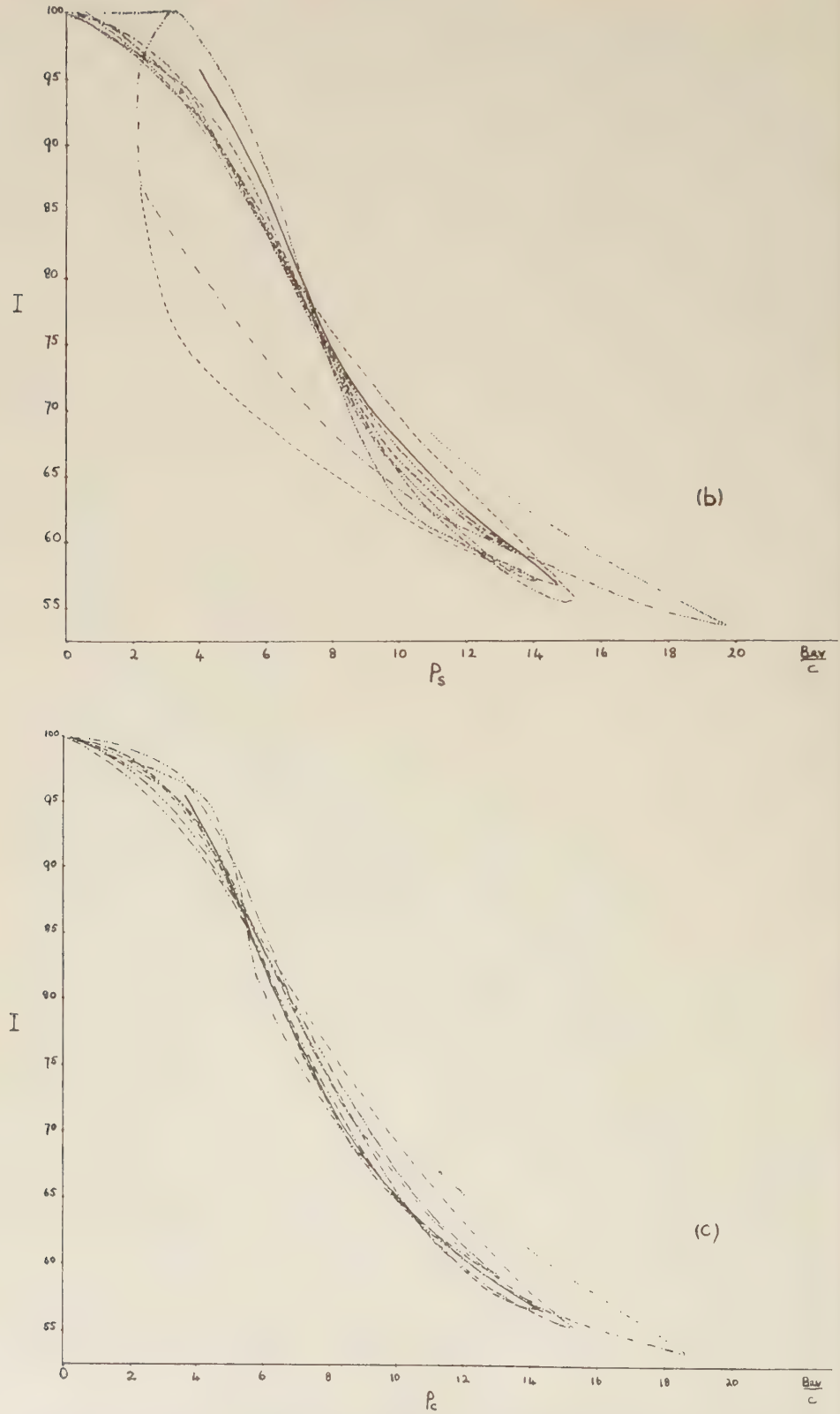


Fig. 3 (*continued*)

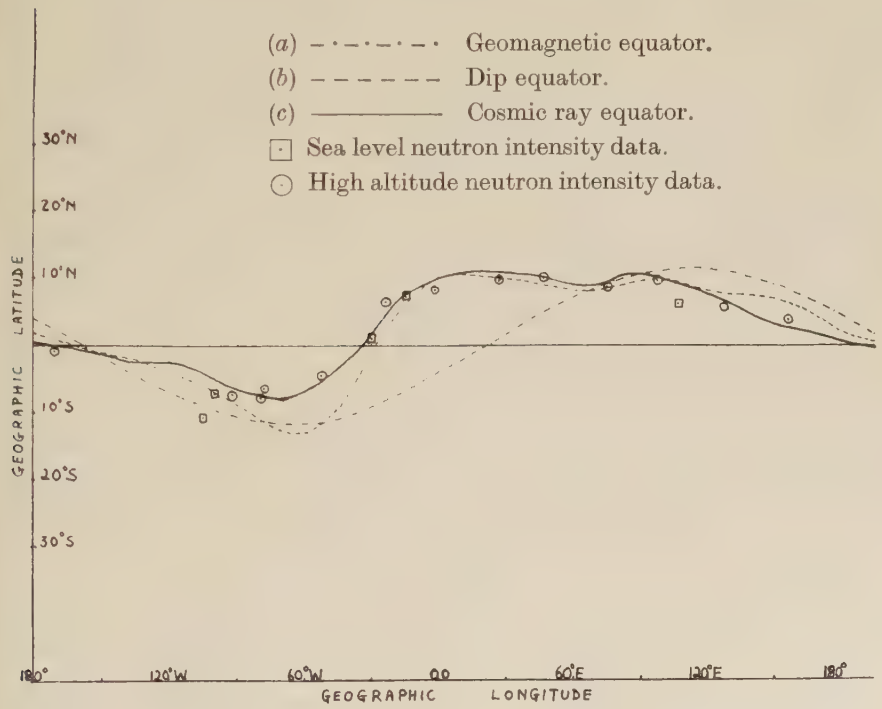


Cosmic ray intensity versus cut-off rigidity on voyages of ' Roxburgh Castle ', ' Atka ', ' Labrador ' and Soya '.

Intensities on ' Roxburgh ' and ' Soya ' have been normalized to same value at Cape Town. Intensities on ' Atka ', ' Labrador ' and ' Soya ' have been normalized to 100 above the '

(2) It has already been shown that the measured positions of minimum cosmic ray intensity lie, to a first approximation, on the dip equator (Rothwell and Quenby 1957). There are some discrepancies, because in the earth's real field, unlike the dipole field, the positions of the maximum horizontal intensity and zero dip do not always coincide. The largest discrepancies between the measured position of minimum cosmic ray intensity and the dip equator occur in the equatorial region of South America, where the greatest distances between the positions of maximum H and zero dip occur. Figure 4 shows the experimentally determined position of minimum intensity reported by Simpson (1957) and others, together with (*a*) the geomagnetic equator; (*b*) the dip equator; and (*c*) the line of minimum cosmic ray intensity computed from eqn. (5) with $x \simeq 0{\cdot}9$: the experimental points all lie very near (*c*).

Fig. 4



Comparison of experimentally determined positions of minimum cosmic ray intensity with (*a*) geomagnetic equator, (*b*) dip equator, and (*c*) cosmic ray equator, determined from eqn. (5).

(3) Table 1, column (*c*) gives the values of α-particle cut-off rigidities calculated from eqn. (5); there is satisfactory agreement with the observed values (although cut-off momenta calculated from the surface field only, p_S, give slightly better agreement with experimental results in several cases).

The expression (5) for the cut-off momenta of cosmic ray particles in the earth's field does, therefore, give satisfactory agreement with experimental results over a wide range of latitudes and longitudes. Hence we conclude that, except in those regions where the earth's real field differs very much from the dipole field, the values of the actual cut-off momenta are close to those obtained by replacing the conventional centre dipole term in the Störmer eqn. (1) by a dipole whose magnitude and direction are determined by the surface field components at the point considered.

Now it is known that the main part of the deflection of a cosmic ray particle occur rather near the earth's surface, at distances of the order of hundreds or at most a few thousand kilometres. Comparison of the dip and the horizontal component of the earth's real field at various distances from the earth's surface and at various latitudes and longitudes (using data compiled by Vestine *et al.* (1947) from the 1945 analysis of the earth's field) with the values of dip and H calculated (*a*) for the dipoles deduced from the surface field components and (*b*) for the conventional centre dipole field, shows that the earth's real field, up to ~ 1000 km above any particular point, is, in fact, better represented by a dipole whose magnitude and direction are deduced from the surface field at that point than by the centre dipole deduced from measurements of the earth's field over the whole of the earth's surface.

It is concluded that discrepancies between simple dipole predictions and experimental observations of cosmic ray intensities and cut-off momenta are due to differences between the earth's field and dipole approximation to it, rather than to distortion of the earth's outer magnetic field by ionized interplanetary matter.

Acknowledgments

The author would like to thank Professor P. M. S. Blackett for helpful advice and criticism, and Dr. W. Webber for his information on measurements of α-particle cut-off momenta.

References

Aly, H. H., and Waddington, C. J., 1957, *Nuovo cim.*, **5**, 1679.
de Marco, A., Milone, A., and Reinharz, M., 1956, *Nuovo cim.*, **3**, 1150.
Fowler, P. H., and Waddington, C. J., 1956, *Phil. Mag.*, **1**, 637.
Fowler, P. H., Waddington, C. J., Freier, P. S., Naugle, J., and Ney, E. P., 1957, *Phil. Mag.*, **2**, 157.
Hopper, V. D., Laby, J. E., and Lim, Y. K., 1958, *Aust. J. Phys.* (to be published).
Kodama, M., and Miyazaki, Y., 1957, *Report of Ionosphere Research in Japan*, **11**, 99.
MacDonald, F. B., 1957, *Phys. Rev.*, **107**, 1386.
Rothwell, P., and Quenby, J., 1957, *Report at Varenna Conference, June* (to be published in *Nuovo cim.*).
Simpson, J. A., 1957, *Report at Varenna Conference, June.*
Simpson, J. A., Fenton, K. B., and Rose, D. C., 1956, *Phys. Rev.*, **102**, 1648.
Vestine, E. H., Laporte, L., Lange, I., Cooper, C., and Hendrix, W. C., 1947, *Carnegie Institute of Washington Publications* 578, 580.
Waddington, C. J., 1956, *Nuovo cim.*, **3**, 930.

Deformation of Thin Films on Solid Substrates†

By D. R. Brame and T. Evans
Tube Investments Research Laboratories, Hinxton Hall, Cambridge

[Received May 20, 1958]

Abstract

Thin films of various face-centred cubic metals in the thickness range of 300 to 700 Å have been oriented on single crystals of silver and palladium. These specimens have then been deformed and the way in which the film accommodates the imposed strain has been determined by examination of the film in the electron microscope after stripping from the substrate. It is considered that the mode of deformation is determined by the ease with which dislocations can be injected into and through the film from the underlying substrate and factors which influence this transfer have been examined.

§ 1. Introduction

Interest was first aroused in the effects of thin films on the mechanical properties of single crystals by the discovery by Roscoe in 1934 that an oxide film, less than 20 atoms in thickness, increased the critical shear stress for slip in cadmium crystals by about 5%. Since that time many workers have confirmed the Roscoe effect with different surface films and substrates. Most of the work has been done with hexagonal close packed crystals such as zinc and cadmium in which there is only one active slip plane (see for instance, Menter and Hall 1950). Andrade and Randall (1948, 1952) found that the presence of hydroxide films on cadmium single crystals completely stopped creep and that twice the stress was required to restart the creep. More recently, Lipsett and King (1957), sputtered thin gold films on cadmium single crystals and the stress at which plastic deformation started was increased by about 6 g wt/mm^2. The increase was found to be independent of film thickness for films in the thickness range of 1500–24 000 Å. The effect with zinc single crystals has been demonstrated using oxide films (Harper and Cottrell 1950), electrodeposited copper (Pickus and Parker 1951, and Gilman 1951), electrodeposited nickel, gold, zinc and silver, and vapour deposited copper and silver (Gilman and Read 1952). With face-centred cubic crystals, very thin oxide films have been shown to increase the critical resolved shear stress of silver single crystals (Andrade and Henderson 1951).

The increase in strength cannot be explained by adding the strength of the thin film to that of the underlying crystal as this would make the strength of the film impossibly large. This implies that the presence of the thin film is modifying the mechanical behaviour of the crystal and three interpretations have been given (Gilman 1955). (1) There is alloying at the interface which strengthens the crystal. (2) The presence of the

† Communicated by J. W. Menter.

surface film inhibits the operation of surface dislocation sources of the Frank–Read type. (3) The surface film prevents exit of dislocations from the surface of the crystal, i.e. slip is inhibited.

Gilman (1955) has investigated the three possibilities and could not detect any significant alloying by spectrographic analyses with crystals which had been strengthened by plating a thin film of copper on their surfaces. He concludes that it is doubtful that alloying has an important effect on the strengthening of the crystals. He tried to differentiate between the second and third possibilities by a comparison of reflected x-rays from the surfaces of deformed plated zinc crystals and deformed clean crystals. He found that the presence of a thin copper film caused more distortion of the surface layers after deformation than for a clean crystal. This supports the hypothesis that the strengthening is due to the film inhibiting the passage of dislocations out of the surface of the crystal since if it inhibited the operation of surface dislocation sources, the distortion would be less for the plated crystal. This explanation is consistent with the work of Barrett (1953) who twisted polycrystalline zinc and steel specimens which had oxide films on them. On removing the oxide layers, the wires twisted a small amount in the original twisting direction. He interpreted this increment in twist as due to the escape from the surface of the dislocations which had piled up under the oxide film as a result of the previous deformation. Measurements of the strengthening of crystals by surface films have been almost exclusively confined to single crystals as the grain boundary strengthening effect in polycrystals is very large and almost completely swamps any strengthening due to surface films.

As a general conclusion from this earlier work it may be stated that the strengthening effect of a thin surface film is associated with its behaviour as a barrier preventing the exit of dislocations from the substrate. Although not explicitly stated, it can be assumed from the conditions of preparation that the films used in previous experiments have been polycrystalline. For the fundamental understanding of the inhibition mechanism, such films are not particularly favourable since there is no control over grain size and relative orientation of the film and substrate. Both these factors are likely to affect the effectiveness of the film to act as a barrier to the exit of dislocations from the substrate. The new experiments described below have been carried out mainly on thin single crystal films prepared by vacuum evaporation on to single crystal substrates so that the film and substrate lattices are in parallel orientation. Under these circumstances the misfit at the boundary between the two lattices can be accommodated by a two dimensional network of dislocations with a spacing determined by the degree of misfit. To examine the parameters which determine the ease of propagation of dislocations through the boundary, a number of face-centred cubic films were oriented on silver single crystals, deformed in tension and the films examined by transmission in the electron microscope after stripping from the silver substrates. The transmission diffraction patterns show the structure of the

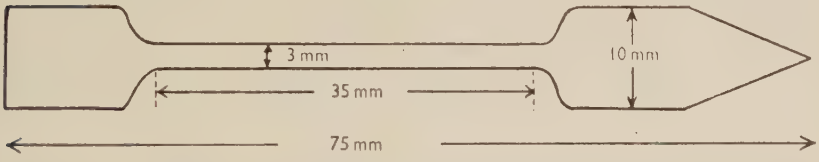
films whilst such details as slip, cracking and dislocation distribution can be obtained by transmission microscopy giving an indication of the way in which the films have accommodated the applied strain when attached to the deformed substrates. Subsidiary experiments were also carried out with polycrystalline films on single crystal silver substrates, oriented films on polycrystalline palladium substrates, and also single crystal films attached to non-crystallographic substrates to determine the types of deformation which the films undergo under various conditions.

§ 2. Experimental Details

Single crystals of silver were produced by a soft mould technique. Tensile specimens of silver with dimensions shown in fig. 1 were cut from a sheet 0·75 mm in thickness. A specimen was packed with alumina in a graphite mould which was passed through a furnace to produce the shaped single crystal. On removal from the mould, the specimen was lightly brushed with a soft brush to remove the alumina and then differentially etched for 30 seconds in the following solution to show the grain structure:

1 pt. of 10 vol. H_2O_2 to 2 pts. of 0·88 HN_4 OH.

Fig. 1



A high yield of crystals was obtained which were single crystal along their gauge lengths and the orientations of these were determined from x-ray back reflection Laue photographs. The specimen was chemically polished by dipping alternately in solutions A and B (Pinner 1953) and then thoroughly washed in distilled water.

A 3 pts. of NaCN (37·5 g/litre)
2 pts. of 10 vol H_2O_2;
B NaCN (37·5 g/litre)

The crystal was placed on a small furnace in a vacuum evaporation plant and the residual surface contamination removed by ionic bombardment at 4 kv with a current density of approximately 60 microamps/cm^2. The method used for the deposition of oriented gold films was an extension of that developed by Pashley (1958) for the production of thin gold films on silver–mica substrates. The silver specimen was maintained at a temperature of 270°c and the gold evaporated from a tungsten filament. The amount of gold evaporated was calculated to correspond to a film thickness of 500 Å, and it was found in practice that the thickness varied in the range of 300–700 Å. This method can be used for the production of thin single crystal films of arbitrary orientations since epitaxial growth occurs

for all surface orientations of the silver substrate. This is useful in such an investigation as silver substrates can be chosen of such orientations that the {111} slip planes in the oriented gold films have large projected areas when viewed by transmission in the electron microscope. The deposition of oriented gold–palladium alloy, platinum and rhodium films of the same nominal thickness as the gold films was done by evaporating from a coiled-coil of tungsten with silver substrate temperature of 250, 350 and 400°c respectively. Deposition was also carried out with the silver substrates at room temperature and oriented single crystals of gold and polycrystalline films of gold palladium alloy, platinum and rhodium were produced by this method. All specimens were deformed by about 15% elongation and the surface film removed by dissolving the substrate in 35% nitric acid. After washing, the film was mounted on a copper grid, dried, and examined by transmission in the Siemens electron microscope (Elmiskop I operating at 80 kv).

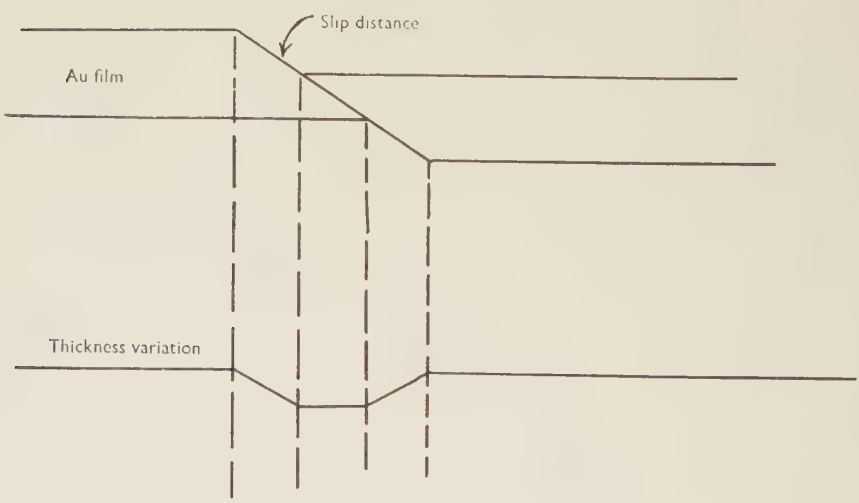
§ 3. Results

3.1. *Single Crystal Films*

3.1.1. *Crystallographic substrates*

Gold films on silver (lattice misfit with silver $< 0{\cdot}2\%$). Examination of an oriented single crystal film after stripping from the silver substrate

Fig. 2



shows that the applied strain is accommodated by a thinning of the film due to slipping on the active (111) planes. The thinning which occurs in a gold film due to slipping is shown diagrammatically in fig. 2. The slipped regions in the electron micrograph are shown as light lines indicating a higher transmission than regions on either side. The gold films remain

coherent and no cracking is observed even after considerable straining (approximately 30% elongation). Figure 3, Pl. 60 shows a gold film which exhibits slipping in two directions due to the substrate orientation being favourably positioned for duplex slip. The contour type of contrast is associated with the buckling of the film and is due to the high intensity of the diffraction when the planes of the film are oriented near the Bragg angle for the incident electrons. When the specimens are deformed in compression instead of tension the slip lines have a reversed contrast showing that the slip is in the opposite direction producing thickening at the slipped regions. Dislocations can be detected in the film by the contrast associated with the stacking fault between the separated partial dislocations (Whelan and Hirsch 1957) which occur in face-centred cubic metals. Figure 4, Pl. 60 shows such a distribution of dislocations and these are considered as dislocations which have been injected into the film from the substrate during deformation and have been retained in the film after stripping.

The behaviour of the oriented gold film when strained on a silver substrate implies that enough dislocations are injected into the film for the applied strain to be accommodated completely in a ductile manner by the formation of slipped regions. If this were not so, the film would crack and this has not been found to occur with gold films. Examination of gold films which have been stripped from the silver substrate without previous deformation shows a complete absence of slipped regions but with a random distribution of dislocations which have been grown into the film. It is not considered that these dislocations are responsible for the slipped regions observed in deformed films since the slip occurs on discrete planes and a general slipping would be expected if in-grown dislocations were responsible. Optical observation of the deformed silver specimen shows no obvious difference between the slip lines on the coated surface and the opposite uncoated surface. This favours the hypothesis that the dislocations responsible for the slipping in the film have been provided by the underlying silver.

In the experiments described here, the substrate temperature during the evaporation of the gold was 270°c but oriented films deposited at room temperature show identical effects after deformation on a silver substrate.

Gold–palladium alloy films with a nominal composition of 40% Au–60% Pd on silver (lattice misfit with silver = 1·9%). Examination of the Au/Pd film after deformation and stripping shows as before that the strain is accommodated in the film by slipping on {111} planes with no evidence for cracking. Figure 5, Pl. 60 shows a Au/Pd film with thinning due to slipping on one set of {111} planes and it is again considered that enough dislocations can be injected into the film to enable it to deform in a ductile manner.

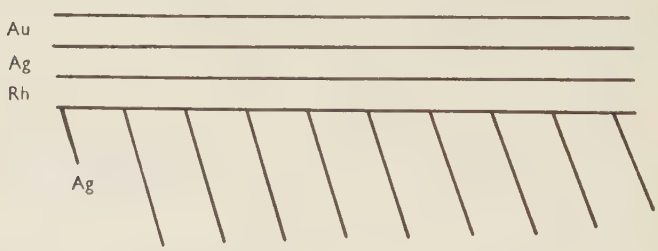
Platinum films on silver (lattice misfit with silver = 4·1%). With oriented films of platinum on substrates of silver slip is again observed

but, in addition, cracks sometimes appear in the general direction of the slip lines. Figure 6, Pl. 61 shows such a crack with the slip lines on either side of it. In this case it appears that the strain is only partly accommodated by slip due to the injection of dislocations and failure presumably occurs because the number injected is too small for the film to deform entirely by slipping. The cracking is envisaged as a brittle failure along a thinned region where the stress is high due to the distortion produced by the piling up of dislocations in the active slip planes of the underlying silver. Figure 7, Pl. 61 shows a platinum film which accommodated the imposed strain entirely by slipping on two slip systems. It is possible that when duplex slip sets in at an early stage in deformation, the sharing of slip on more than one set of slip planes makes the passage of dislocations from the substrate into the film more favourable than when slip on a single set of planes occurs extensively. This would then make the passage of dislocations dependent upon the orientation of the system.

Rhodium films on silver (lattice misfit with silver = 7·37%). With rhodium films on silver substrates, extensive cracking occurs with a slight amount of slipping. Figure 8, Pl. 61 shows such an area with a slight amount of slipping and fig. 9, Pl. 61 illustrates cracking without the presence of slip. As with platinum, it is considered that not enough dislocations can be injected into the rhodium film to accommodate the applied strain and cracking occurs along the highly stressed regions above the active slip planes in the silver substrate.

Barrier effect of rhodium films. The experiment described above suggests that it is difficult for dislocations to pass from a silver substrate into and through a rhodium film. In this sense, the interface may be considered

Fig. 10

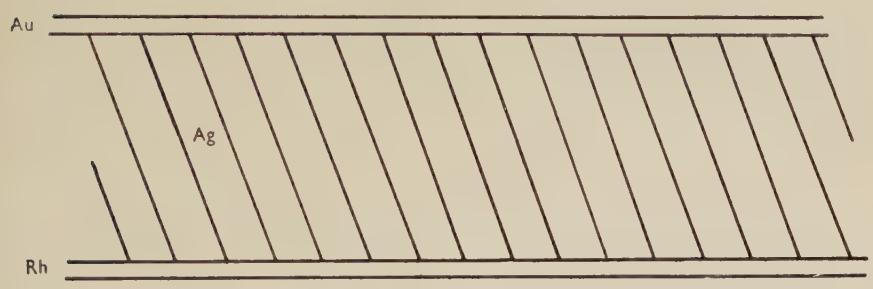


as a barrier to the passage of dislocations and a multilayer of thin films has been grown on a bulk silver substrate to test this. Thin oriented layers of rhodium, silver and gold were successively evaporated on to a bulk silver single crystal substrate producing a specimen as shown in fig. 10. This was then deformed in tension and the gold film detached by dissolving away the silver. Examination of the gold film showed that it had cracked in an apparently brittle fashion, as shown in figure 11, Pl. 62. It seems probable that the rhodium film is acting as a barrier to the passage of dislocations from the silver substrate into the multilayer, and the gold,

being on the outside, accommodates the imposed strain by cracking because not enough dislocations can be injected into it to enable slip to take place.

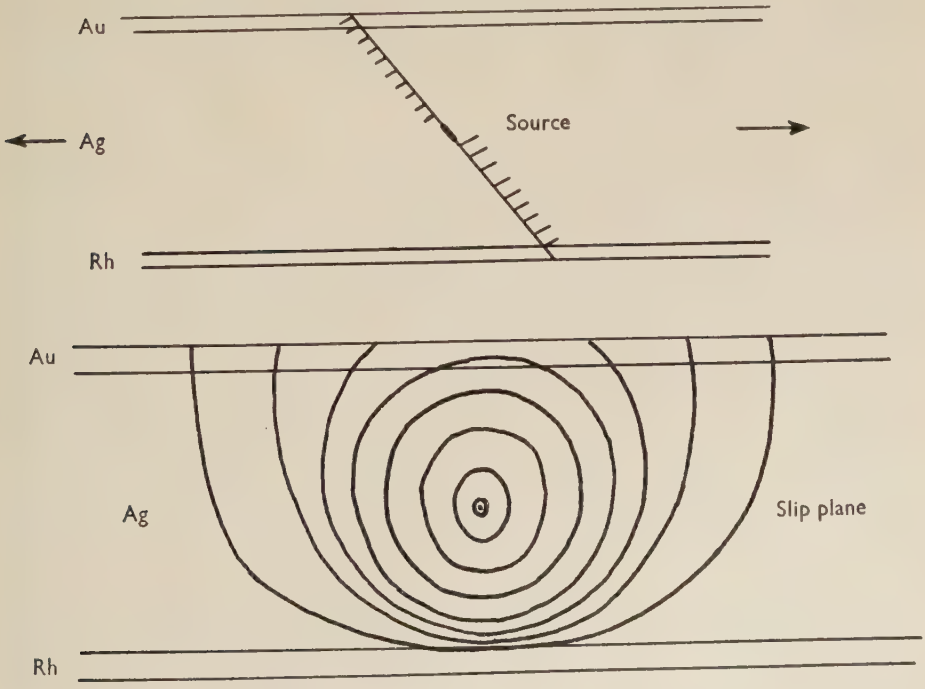
The barrier effect of the rhodium film-silver substrate interface was demonstrated in another way. An oriented single crystal film of rhodium was evaporated on one side of a single crystal tensile specimen of silver and an oriented film of gold evaporated on the other side, as shown in fig. 12. This specimen was deformed in tension and the gold and rhodium

Fig. 12



films examined subsequently. The rhodium film had cracked as before and the gold film showed a far higher density of injected dislocation groups than was noted when the rhodium film was absent, as shown in

Fig. 15



figs. 13 and 14, Pl. 62. This is attributed to the fact that the rhodium film, in places where it is not cracked, acts as a strong barrier to the passage of dislocations which pile up against the rhodium film boundary. This produces a strong back stress which stops active sources from operating. The gold–silver interface allows the passage of dislocations through it and thus a distorted configuration of dislocation loops is produced on the active slip planes as shown schematically in fig. 15. The rhodium effectively pins the loops in the specimen and the gold film can trap segments of the loops which have passed through the gold film but not transferred through the rhodium film. In addition, the hardening due to the presence of the rhodium film causes more dislocation sources to become operative in the silver substrate. This will occur in the substrate slip planes over which the rhodium film has not cracked as cracking would enable the dislocations to escape easily from the silver substrate on the rhodium side.

Gold films on palladium (lattice misfit = 5%). Large grained polycrystal specimens were used as substrates on account of the difficulties of making single crystals with the material available. Comparison with experiments using similar silver substrates indicates that provided grain boundary regions are ignored the behaviour of the film on any particular grain may be taken as typical of a single crystal in the corresponding orientation.

It was found that the strain in an oriented gold film on a palladium substrate is accommodated entirely by slipping.

Rhodium films on palladium (lattice misfit = 2·26%). Again the strain is accommodated almost entirely by slipping although occasional cracking does sometimes occur. The significance of these experiments on palladium substrates in relation to those on silver is discussed below.

3.1.2. *Non-crystallographic substrates*

The thinning of the gold films on silver substrates by slip during deformation suggests that the dislocations required for this mode of deformation arise from the active slip planes in the crystallographic substrate. To test this, a single crystal gold film was attached to a rubber balloon which served as a non-crystallographic substrate. Optical observation of the gold film when the balloon was inflated showed that the film progressively cracked into smaller pieces. This brittle cracking is attributed to the absence of injected dislocations from the substrate to accommodate the imposed strain. Owing to the complex stressing conditions with balloons similar experiments were carried out on a film attached to a tensile specimen of a polymer when deformation again produced a brittle type of cracking in the film with the crack direction normal to the applied stress.

3.2. *Polycrystalline Films*

By deposition with the substrate at room temperature, unoriented polycrystalline films of gold-palladium alloy, platinum and rhodium with

a grain size of less than 100 A have been formed on single crystal substrates of silver. The film and substrate were again deformed in tension and the film examined in the electron microscope after stripping from the silver substrate. In these cases, the strain was accommodated in the film by cracking along the directions of the slip lines on the surface of the underlying silver substrate. Figure 16, Pl. 62 shows a typical region of cracking in one direction. These films are considered to be the type which have been used in the published work on the strengthening effect of films on single crystals.

§ 4. Discussion

It is first necessary to consider the origin of the dislocations which permit a surface film to deform by slip when it is oriented on a substrate which is deformed plastically. A number of alternatives are possible.

(1) Deformation by virtue of the movement of the dislocations grown into the film.

(2) Activation of dislocation sources in the film.

(3) Deformation by virtue of dislocation sources operating in the bulk substrate which emit dislocations able to pass into and through the film.

(1) The experiment with a Au/Ag Rh multilayer on silver indicates that extensive plastic deformation of a gold film by the movement of dislocations is not likely. This is consistent with observations by Pashley (1958) that unattached thin single crystal metal films of this thickness (~ 500 Å) are inherently brittle. Furthermore the visual evidence from experiments where the film is ductile (e.g. Au on Ag) shows that the deformation in the film occurs by extensive glide on single or closely spaced slip planes. The grown in dislocations distributed at random throughout the film would not produce plastic deformation of this type even if they were able to move.

(2) Such regions of extensive glide could be produced by the activation of sources within the film. Again the experiment with the Au/Ag/Rh multilayer on silver shows that this is extremely unlikely. There is the additional possibility that dislocation sources of the type observed by Whelan, Hirsch, Horne and Bollmann (1957) in stainless steel could operate. They found that dislocations can be nucleated at a comparatively low stress in a wedge shaped film such as might occur near a hole or an edge. Although there are sometimes holes in the films they are much too infrequent to account for the deformation, and slipped regions in the film are never observed to be specially associated with such holes as are present. A wedge shaped film may be formed at the edges of a specimen as shown at A and B in fig. 17. In order to eliminate this as a possible source of dislocations a silver specimen was coated all round with a gold film of uniform thickness as shown in fig. 18. No difference was observed in the characteristics of the deformation of the film and it is concluded that possible wedges at the edges are not responsible for nucleating the dislocations which result in the ductile behaviour of the gold film.

(3) All the experimental evidence is consistent with the assumption that the behaviour of the film is determined by its reaction to dislocations arising from sources activated in the interior of the substrate. These produce dislocation loops on the active slip planes and the mechanical response of the film depends upon the ability of these dislocations to pass into and through the film. Factors which affect the ease of passage of dislocations into the film include:

(*a*) alloying between the substrate and the film,

(*b*) the structure of the film itself,

(*c*) the difference in elastic moduli of the substrate and the thin film, and

(*d*) the degree of misfit between the substrate and the film.

Fig. 17

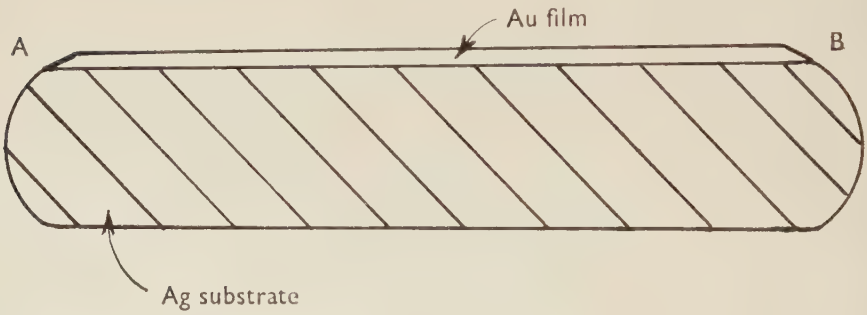
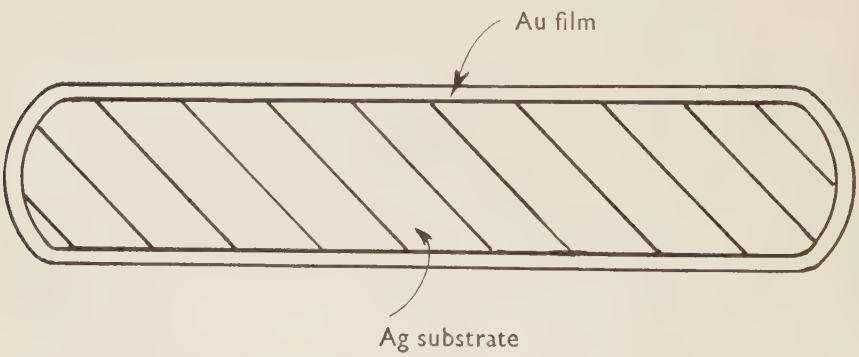


Fig. 18



(*a*) The extent to which alloying occurs during the growth of these films is at present uncertain and the problem is being investigated in the laboratory. However, it has been found that oriented gold films on silver substrates show the same type of deformation whether they have been grown with the substrate at room temperature or at 270°c. This suggests that in this system at least, the effects of interfacial alloying are small as

presumably less alloying would occur during deposition with the substrate at room temperature.

(*b*) A perfect oriented single crystal film has its {111} slip planes parallel to those in the substrate and a dislocation can pass from the substrate slip plane into the film. Thus no misfit is left behind at the boundary during the transfer due to the slip plane in the film being misoriented with respect to the substrate slip plane. On the other hand with a polycrystalline film on a single crystal substrate it is very difficult for a dislocation to be transferred because the slip planes in the grains of the film are randomly oriented with respect to one another and also to the active slip planes in the substrate. Thus a polycrystalline thin film on a single crystal substrate accommodates the imposed strain by cracking in the directions of the slip traces in the underlying silver. These are presumably where there are stress concentrations due to the distortion produced in the substrate by piled up dislocations below the film-substrate boundary.

(*c*) Head (1953) has analysed the image forces arising when a screw dislocation in a substrate approaches a surface to which is attached a thin film of different elastic modulus. If the elastic modulus of the film is greater than that of the substrate, a dislocation approaching the surface experiences first an attraction up to a critical distance and thereafter a repulsion from the surface. The equilibrium distance of the dislocation from the boundary depends upon the relative elastic moduli and the film thickness. The critical distance varies from less than 50 Å in the case of a 400 Å gold film on silver, half the film thickness for platinum on silver and greater than the film thickness in the case of rhodium film on silver. Head's analysis is concerned with a screw dislocation but he considers that the same conditions hold for an edge. The shear moduli of the metals used in these experiments are given in the table.

Metal of film	Degree of misfit with substrate	Shear modulus (Million p.s.i.)
Silver Substrate	—	3·9
Gold	<0·2%	4·1–5·6
Gold/Palladium	1·9%	—
Palladium	5·0%	6·7
Platinum	4·1%	8·5–10
Rhodium	7·37%	16–20
Palladium Substrate	—	6·7
Gold	5%	4·1–5·6
Rhodium	2·26%	16–20

The modulus of the film is always greater than that of a silver substrate. We may therefore expect that in all these cases the Head effect will tend to inhibit the exit of dislocations from the substrate with a repulsion which increases as one progresses from gold to rhodium films. Thus,

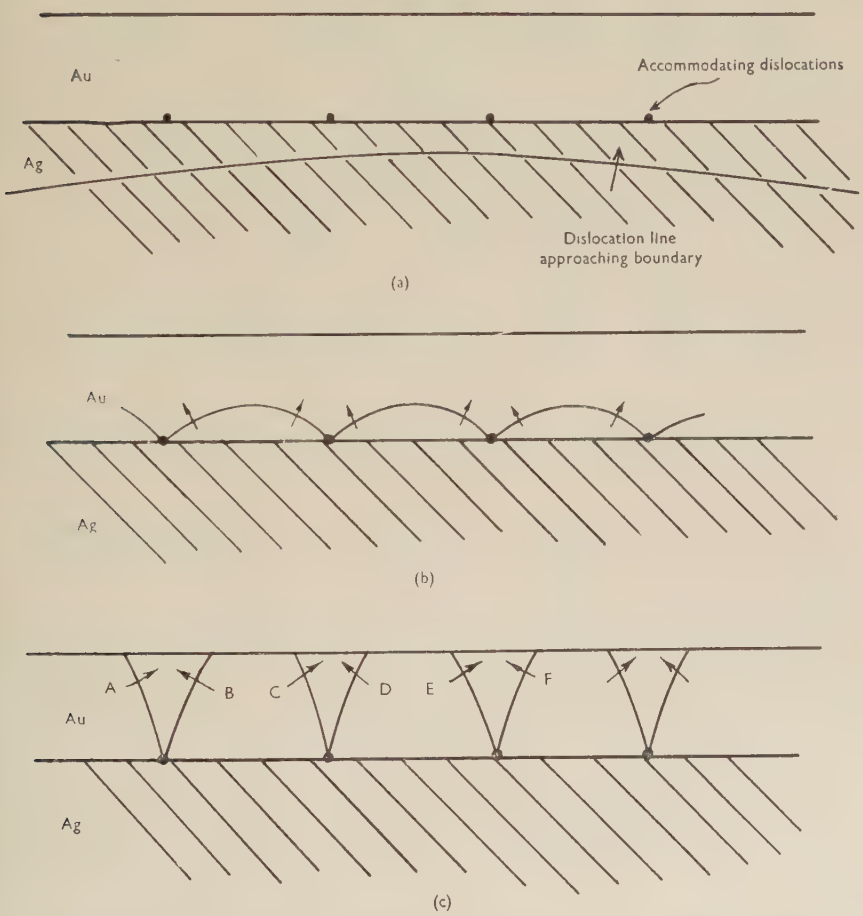
with silver substrates and the films which have been used, as the degree of misfit between the two lattices increases so does the repulsion due to the relative elastic moduli. To try to differentiate between the two effects, gold and rhodium films in turn were deposited on palladium substrates. In the gold case, the modulus of the film is less than that of the substrate which, according to Head's analysis, will result in there being an attractive force on a dislocation approaching the surface. As the strain in the film is entirely accommodated by slip, although the lattice misfit is 5%, it is concluded that the ratio of the shear moduli has some effect on the passage of dislocations from the substrate into the film. In the case of rhodium on palladium there is a large difference in the moduli with the film having the higher modulus. This time, slip again occurs in the film with very occasional cracking with a lattice misfit of 2·26% which again indicates that the ratio of the shear moduli is having some effect.

(*d*) The main influence in determining the transfer of dislocations from the substrate to an oriented film is thought to be the difference in lattice parameter between them. Firstly, the barrier effect of the accommodating network of dislocations between them (van der Merwe 1950) must be considered. The dislocations of this network are mobile only in the boundary and their spacing depends upon the misfit between the two lattices and the orientation of the interface. On a (111) interface the network has trigonal symmetry, whereas on (110) interfaces it is rectangular and square on (100) interfaces. Gold, gold–palladium alloy, platinum and rhodium have a lattice misfit with respect to silver of $<0{\cdot}2\%$, $1{\cdot}9\%$, $4{\cdot}1\%$ and $7{\cdot}37\%$ respectively, and the spacing of the accommodating network on a (111) interface would vary from the order of 2000 Å in the case of gold to 32 Å in the case of rhodium. With a gold film on a silver substrate, the hindrance to the passage of dislocations through a network of 2000 Å spacing is expected to be small. For instance, in the particularly simple case where the accommodating dislocations being cut have a Burgers vector parallel to the mobile dislocation as it reaches the boundary, transference is possible by the following mechanism.

In fig. 19, (*a*) shows the mobile dislocation line in the silver approaching the silver-gold boundary. At (*b*) the mobile dislocation is pinned at the accommodating dislocations and the segments bow out between the pinning positions. At (*c*) the segments have cut the free surface and move towards one another and A combines with B, C with D, and E with F. This results in a slipped region in the gold film itself by the effective transference of the mobile dislocation from the silver substrate through the gold film. The stress required by such a cutting mechanism is extremely small and is $\mu_1\mathbf{b}/2500$ where μ_1 = shear modulus of gold and $\mathbf{b}$ = Burgers vector of the dislocation in the gold. This assumes a film thickness of 500 Å and a network spacing of 2000 Å. This is a particularly simple case and does not involve jog formation but it does serve to illustrate the weakness of the accommodating network in the case of a gold film on a silver substrate. In the general case, the Burgers vector of the dislocations being cut will

not be parallel to the cutting dislocation and jog formation will occur. As the misfit increases and the network spacing consequently decreases, the stress required for a dislocation to cut through will increase. Cottrell (1953) has pointed out that the stress required for a dislocation to cut through a row of static dislocations is inversely proportional to the spacing between the static dislocations and it again appears that as films of gold, gold palladium, platinum and rhodium are used in turn on silver substrates it will be progressively more difficult for the mobile dislocation to cut through the interfacial network owing to the decreasing network spacing.

Fig. 19



The actual transfer of a dislocation from a single crystal silver substrate to an oriented film of a different material will now be considered as this may contribute a strong barrier to the passage of dislocations. When a dislocation passes from a substrate with a particular lattice parameter into an oriented film with a different lattice parameter, the Burgers vector

changes in crossing the boundary and a portion of the Burgers vector is left behind at the interface. Consider a dislocation of Burgers vector $\mathbf{b}_1$ in the substrate. On passing into the film the Burgers vector changes to $\mathbf{b}_2$ and when a silver substrate is used $|\mathbf{b}_2| < |\mathbf{b}_1|$. After n dislocations have passed the accumulated Burgers vector left behind at the boundary is $n(\mathbf{b}_1 - \mathbf{b}_2)$. This is illustrated in fig. 20. The energy at region CD can increase in the slip plane until $n(\mathbf{b}_1 - \mathbf{b}_2) = \mathbf{b}_2$ (or more probably $> \frac{1}{2}\mathbf{b}_2$) which would create a dislocation in the slip plane at CD. This dislocation can then pass into and through the film and the situation shown in fig. 21

Fig. 20

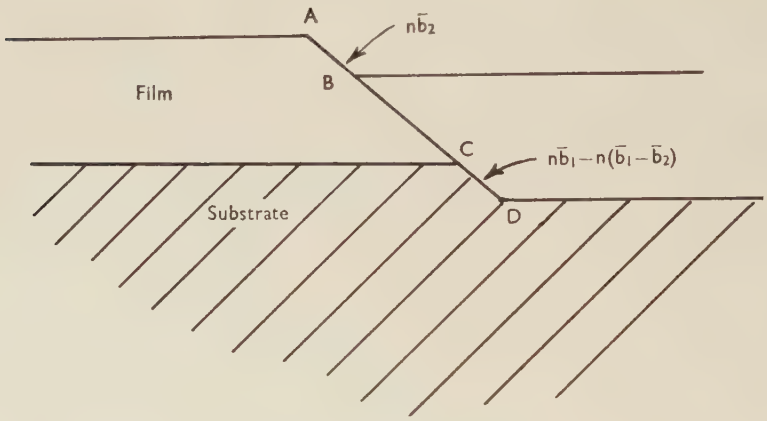
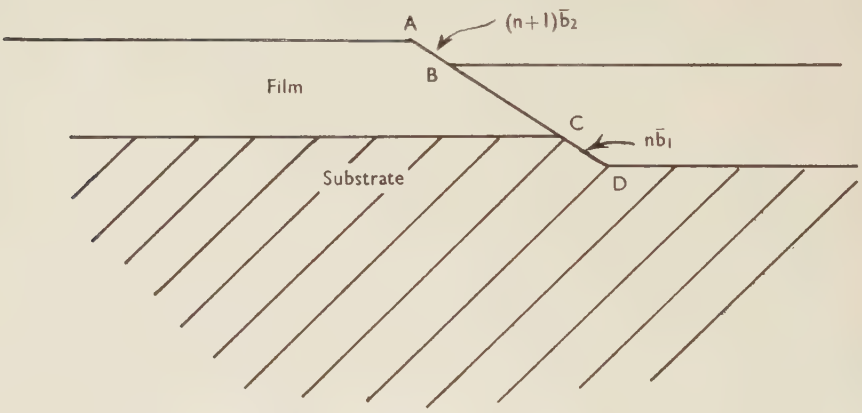


Fig. 21



would be realized. Then the process can be repeated by a further accumulation of Burgers vector at CD with a further creation of dislocations which can move through the film. The slip plane at CD is a region of misfit between the substrate and film lattices which can be accommodated by the formation of an accommodating dislocation network at the CD interface. It is not possible for the Burgers vector residues left behind at CD

by the passage of dislocations from the substrate to the film to build up the accommodating network at CD and it seems essential for two types of dislocation to be formed at the interface at CD; one to relieve the increasing strain due to the Burgers vector residues left behind at the interface and the other to relieve the localized strains due to the mismatch between the two lattices on either side of the CD portion of the slip plane. When $|\mathbf{b}_2| > |\mathbf{b}_1|$ as in the case of an oriented gold film on a palladium substrate, the same type of mechanism occurs at CD except that the extra dislocation created by an accumulation of Burgers vector residues would have the opposite sign to those moving on the active slip plane of the substrate. This dislocation then combines with one of the mobile dislocations and they annihilate so that $(n+1)$ dislocations move up the substrate slip plane to CD and n dislocations emerge at the free surface AB. An accommodating network of dislocations would again be necessary in the portion of the slip plane at CD to relieve long range stresses due to the lattice mismatch on either side. The accumulation of the Burgers vector residues occurs in the active slip plane as does the accommodating network at the region CD. These dislocations can act as a strong barrier to the passage of dislocations on an active slip plane as they lie in the active slip plane and the effect increases as the number of dislocations passes from the substrate into the film. Again, the rate of increase of such a hardening process is directly related to the degree of misfit of the two lattices since both the number of dislocations n required before a dislocation can be created at CD and the spacing of the accommodating network decreases as the mismatch increases. The portion of slip plane CD will be so small that the normal network of accommodating dislocations for a (111) interface with trigonal symmetry will probably not be reached but some system seems to be necessary.

Experimentally, it seems that, in the case of an oriented gold film on a silver substrate at least, the active slip planes become inoperative before the stage is reached where a dislocation is created at the interface due to the accumulation of Burgers vector residues. This is because it is observed that the gold films even as thin as 100 Å do not break due to sliding off on the slip planes. This means either that the normal slip distance of silver crystals is less than the order of 100 Å anyway or the barrier at the interface due to the presence of an oriented gold film makes the active dislocation sources inoperative at an earlier stage and the slip steps are smaller.

In the case of oriented thin films of the other materials used, it is possible for the stage to be reached where the accommodating network in the active slip plane is formed at the interface and where the residue of Burgers vectors creates an extra dislocation.

As a general conclusion to this work it is considered that the mode of deformation of an oriented film on a substrate depends upon the ability of dislocations to be injected into the film from the underlying substrate. If enough dislocations can be injected, the film will deform in a ductile manner, if not, the film will crack. The experimental evidence indicates that the

controlling factors are the degree of misfit and the relative elastic moduli of the film and substrate. Although the former is predominant no quantitative criterion for the maximum permissible misfit to obtain ductility can be given since the relative moduli of film and substrate undoubtedly influence the behaviour of the film.

Acknowledgments

The authors wish to thank Dr. J. W. Menter for his guidance throughout this work and Drs. B. A. Bilby, A. J. Forty and D. W. Pashley for useful comments. This paper is published by permission of the Chairman of Tube Investments Limited.

References

Andrade, E. N. da C., and Henderson, C., 1951, *Phil. Trans.*, **244,** 177.

Andrade, E. N. da C., and Randall, R. F. V., 1948, *Nature, Lond.*, **162,** 890 ; 1952, *Proc. phys. Soc. Lond.* B, **65,** 445.

Barrett, C. S., 1953, *Acta Met.*, **1,** 2.

Cottrell, A. H., 1953, *Dislocations and plastic flow in crystals* (Oxford : Clarendon Press).

Gilman, J. J., 1951, *Trans. Amer. Inst. min. (metall.) Engrs*, **191,** 1148 (*J. Metals, N.Y.*, **3**) ; 1955, *A.S.T.M. Special Technical Publication* No. 171, 3.

Gilman, J. J., and Read, T. A., 1952, *Trans. Amer. Inst. min. (metall.) Engrs*, **194** (*J. Metals, N.Y.*, **4**).

Harper, S., and Cottrell, A. H., 1950, *Proc. phys. Soc. Lond.* B, **63,** 331.

Head, A. K., 1953, *Phil. Mag.*, **44,** 92.

Lipsett, F. R., and King, R., 1957, *Proc. phys. Soc. Lond.* B, **70,** 608.

Menter, J. W., and Hall, E. O., 1950, *Nature, Lond.*, **165,** 611.

Pashley, D. W., 1958 (to be published).

Pickus, M. R., and Parker, E. R., 1951, *Trans. Amer. Inst. min. (metall.) Engrs*, **191,** 792 (*J. Metals, N.Y.*, **3**).

Pinner, R., 1953, *Electroplating*, **6,** 401.

Roscoe, R., 1934, *Nature, Lond.*, **133,** 912.

van der Merwe, J. H., 1950, *Proc. phys. Soc. Lond.* A, **63,** 616.

Whelan, M. J., and Hirsch, P. B., 1957, *Phil. Mag.*, **2,** 1303.

Whelan, M. J., Hirsch, P. B., Horne, R. W., and Bollmann, W., 1957, *Proc. roy. Soc.* A, **240,** 524.

Anharmonic Effects in the Theory of Solid Argon†

By I. J. Zucker
Wheatstone Physics Laboratory, King's College, Strand, W.C.2‡

[Received April 19, 1958]

Abstract

The thermodynamic properties of solid argon are evaluated theoretically using the Debye quasi-harmonic theory and an Einstein theory modified to include anharmonic effects. The anharmonic theory gives better agreement with experiment except for the specific heat at low temperatures.

§ 1. Introduction

An excellent review of the theory and properties of solid argon has been given by Dobbs and Jones (1957), and the purpose of this communication is to discuss in more detail some of the theoretical results given in that review.

The inert gas crystals of which solid argon is a member have been the study of many theoretical investigations because of their simplicity. Recent experimental measurements of the properties of solid argon by Stewart (1955, 1956), Barker and Dobbs (1955) and Dobbs *et al.* (1956) have provided useful material for comparing various models and theories concerning crystal lattices. Einstein (1907, 1911 a, b) put forward the first reasonably successful approximation to the thermodynamic behaviour of a crystal. It was assumed that the constituent particles of a crystal all vibrated independently of one another about their mean rest positions with the same frequency. Debye (1912) improved this model by taking into account in an approximate way the interdependence of the particles. A crystal was replaced in this approximation by an elastic continuum and the normal modes of vibration of the latter were assumed to represent the normal modes of the crystal lattice. More refined attacks deriving from the original work of Born and von Karman (1912, 1913) were based on solving the dynamical many body problem of a lattice. All these approaches had one feature in common. This was that the particles of a crystal lattice were constrained to their equilibrium positions by harmonic forces. This was equivalent to only considering terms up to the second order in the expansion of the potential energy as a Taylor series in the small displacements of the particles from their equilibrium positions. Although purely harmonic theories cannot account for anharmonic properties, it is possible to modify these theories by allowing parameters

† Communicated by the Author.

‡ Now at Research Laboratories of the General Electric Company Limited, Wembley, England.

such as the Debye value to vary with volume and this can take account of properties such as thermal expansion and compressibility. The term quasi-harmonic will be used in reference to such theories.

Quasi-harmonic theories are found to be moderately successful in predicting the properties of solid argon at low temperatures, but are much less successful at temperatures near the melting point. This suggests that some account of anharmonic terms in the potential energy should be made. The difficulty of extending the more refined theories of lattice vibrations to include anharmonic terms is formidable. Though Born (1951) and Hooton (1955 a, b) have made progress in this direction it is not possible to apply their results without making other assumptions. But for the Einstein model, Henkel (1955) has shown in a particular case how higher order terms in the potential energy may be evaluated. It is well known that thermodynamic properties of crystals evaluated using different harmonic theories give almost the same results. Thus although the Einstein model is rather crude, calculations made with it including anharmonic terms should indicate differences between harmonic and anharmonic theories. In this paper a comparison is made between results calculated by using a quasi-Debye model and an Einstein model including an anharmonic term. Henkel has already made some calculations but these will be repeated and amplified using what is believed to be a more accurate representation of $\phi(r)$ the potential energy between a pair of argon atoms. This is

$$\phi(r) = \frac{\lambda}{r^{12}} - \frac{\mu}{r^6}$$

$$\lambda = 1{\cdot}63 \times 10^{-7}\,\text{Å}^{12}\,\text{ergs},\ \mu = 1{\cdot}05 \times 10^{-10}\,\text{Å}^{6}\,\text{ergs} \quad . \quad . \quad . \quad (1)$$

is determined from solid state data alone, the heat of sublimation and lattice constant at absolute zero being the information used. The method of obtaining and the reasons for this choice are given elsewhere (Domb and Zucker 1956, Zucker 1956).

§ 2. Theory

In the following sections the letter D or H in brackets following a symbol referring to a thermodynamic quantity implies that the latter is found by use of the quasi-Debye or Henkel theory respectively.

The free energy of a Debye crystal is given by

$$F(\mathrm{D}) = \tfrac{1}{2}N\textstyle\sum_r \phi(r) + \frac{9}{8}R\theta_\mathrm{D} + 9R\frac{T^4}{\theta_\mathrm{D}^3}\int_0^{\theta_\mathrm{D}/\mathrm{T}} \log(1-e^{-x})x^2\,dx. \quad . \quad (2)$$

$\sum_r$ indicates summation over all lattice points of a crystal. The first term of (2) represents the static lattice energy, the second term the zero-point energy and the last term the thermal energy. θ_D is defined as $h\nu_\mathrm{D}/k$† where ν_D is the maximum frequency of the elastic continuum equivalent to the lattice. Domb and Salter (1952) have given a simple

† h is Planck's constant and k is Boltzmann's constant.

method of finding θ_D as a function of volume. They relate ν_D^2 to the sum of the squares of the frequencies of the normal modes of a crystal. θ_D is then found in terms of the force constants which are in turn given in terms of $\phi(r)$. Thus

$$\left.\begin{aligned}\theta_D^2 &= \frac{5h^2N^2}{36k^2\pi^2M}\left[\sum_r \phi''(r)+\frac{2}{r}\phi'(r)\right] \\ &= \frac{5h^2N^2}{36k^2\pi^2M}\left[132\,\frac{\lambda}{r^{14}}c_{14}-30\,\frac{\mu}{r^8}c_8\right] \\ c_{14} &= 12{\cdot}06 \quad c_8 = 12{\cdot}80\end{aligned}\right\} . \quad . \quad . \quad (3)$$

The C_n are numbers obtained by summing inverse powers over all lattice points of a face-centred cube in terms of the distance between nearest neighbours. These were originally evaluated by Lennard-Jones and Ingham (1925). The method of obtaining θ_D here is simpler and probably more accurate than that proposed by Herzfeld and Goeppert-Mayer (1934). It also provides a check on $\phi(r)$. Putting in the value of the nearest neighbour distance at the absolute zero one finds $\theta_D = 81$, in good agreement with the experimental value of 80.

From (3) all thermodynamic quantities may be determined employing the usual thermodynamic relations.

Henkel extended the Einstein theory by finding the potential energy of an atom with respect to all the others up to the fourth order term. He did this for a particular form of $\phi(r)$ for a face-centred cubic lattice. This may be generalized to any central additive potential and to any lattice. The potential energy of an atom is found to be

$$\begin{aligned}V &= P_0+P_2(x^2+y^2+z^2)+P_4(x^4+y^4+z^4) \\ P_0 &= \tfrac{1}{2}\sum_r \phi(r), \quad P_2\dagger = \frac{1}{6}\sum_r \phi''(r)+\frac{2}{r}\phi'(r), \\ P_4 &= \frac{1}{72}\sum_r \phi^{iv}(r)+\frac{4}{r}\phi'''(r). \quad . \quad . \quad . \quad . \quad . \quad (4)\end{aligned}$$

The energy levels for an atom with potential energy given by (4) are easily found since the Schrödinger equation is separable into three equations each of the form

$$\left(-\frac{h^2}{8\pi^2 m}\frac{\partial^2}{\partial x^2}+P_2x^2+P_4x^4\right)\Psi = \epsilon\Psi. \quad . \quad . \quad . \quad . \quad (5)$$

† It will be observed that P_2 appears in the formula for θ_D. Indeed $6P_2$ is the trace of the $3N \times 3N$ matrix obtained by lattice dynamics, the eigenvalues of which determine the squares of the frequencies of the normal modes of the lattice. The difference between the ordinary Einstein and Debye theories is due to the different ways of evaluating the respective characteristic temperatures by relating the respective frequency distributions to $6P_2$. The difference between θ_E and θ_D found in this manner is only a numerical factor; in fact $\theta_D^2 = 5/3\theta_E^2$. Since the volume dependence is obviously the same the only significant differences between the two theories employing characteristic temperatures calculated as indicated is in the specific heat at low temperatures. Here the usual differences between the Einstein and Debye functions exist.

This may be solved simply by treating the $P_4 x^4$ term as a perturbation whence the free energy per mode of a crystal in the Henkel approximation becomes

$$F(\mathrm{H}) = \tfrac{1}{2}N\sum_r \phi(r) + \frac{3}{2}NW - 3RT\log\sum_{n=0}^{\infty}\exp\left(-\frac{1}{kT}\right)(nW + n^2Y),$$

$$W = \frac{h}{2\pi}N\left(\frac{2P_2}{M}\right)^{1/2}, \quad Y = \frac{3h^2N^2}{16\pi^2M}\frac{P_4}{P_2}. \qquad (6)$$

It will be observed that when P_4 is zero the result becomes that for the free energy of an ordinary Einstein crystal, and W/k may be identified as the Einstein characteristic temperature.

§ 3. Evaluation of the Thermodynamic Properties of Solid Argon

3.1. *Density as a Function of Temperature*

Finding the density is equivalent to finding the equilibrium molar volume under zero pressure at various temperatures. Thus $P = -(\partial F/\partial V)_T$ was found from eqns. (2) and (6). The value of V which made P zero was found in both cases at 10° intervals from 0° to 80°K.

3.2. *Isothermal Compressibility K_T at $P = 0$*

K_T was found using the relation $K_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$. The values of V found already were used.

3.3. *Expansivity α, as a Function of Temperature*

α is equal to $1/V\ (\partial V/\partial T)_P$ and this is equivalent to $K_T\ (\partial P/\partial T)_V$. K_T having been evaluated, $(\partial P/\partial T)_V$ was calculated, and hence α.

3.4. *Specific Heats as Functions of Temperature*

C_V is given by $-T\left(\frac{\partial^2 F}{\partial T^2}\right)_V$ and C_P by

$$C_P = C_V + \frac{\alpha^2 VT}{K_T}.$$

C_P was evaluated for both theories using the respective values of C_V, α, V and K_T already found.

K_T was also found as a function of pressure at five different temperatures but only the Henkel equations were employed in this case. All these results have been illustrated graphically in figs. 1–6 and tabulated below.

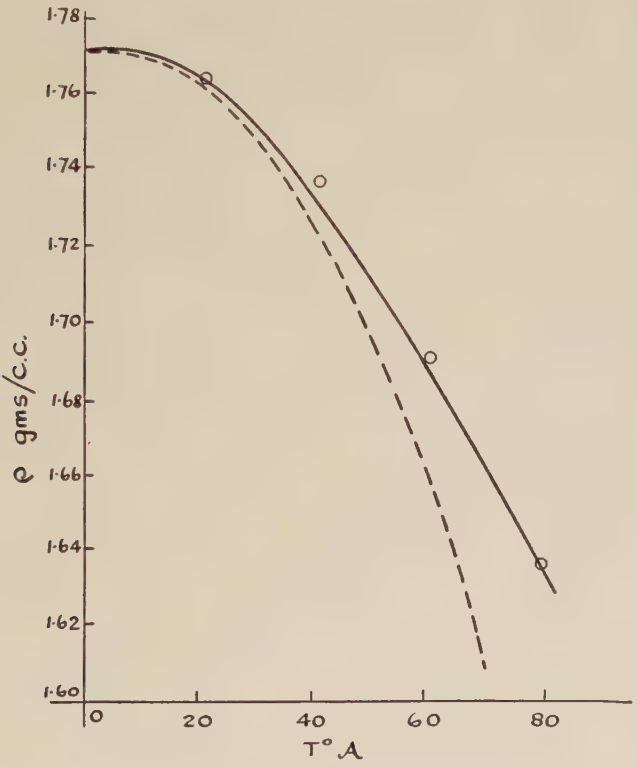
§ 4. Discussion of Results

4.1. *ρ vs T*

Since ρ at $T = 0$ is one of the parameters used in determining $\phi(r)$ it is not surprising that the theoretical and experimental values agree at $T = 0°$. This point is in fact the only fixed point in all the calculations described

above. In fig. 1 it is seen that the differences between ρ(D) and ρ(H) are small, but increase rapidly with temperature. ρ(H) is undoubtedly in better agreement with the experimental values of Dobbs *et al.* (1956). The ρ(D) are all too small and in fact it was found that at 80°K there was no volume at which the crystal was in equilibrium when the pressure is

Fig. 1



Density against temperature. ———Henkel theory. — — — — —Debye theory. O Experimental Results—Dobbs *et al.*

zero. The value of ρ(D) given is that which makes the pressure a minimum. It may be deduced from this fact that a harmonic crystal is more expanded and less stable than an anharmonic crystal. The reason is that for a given energy the amplitude of a harmonic oscillator is greater than for an anharmonic oscillator.

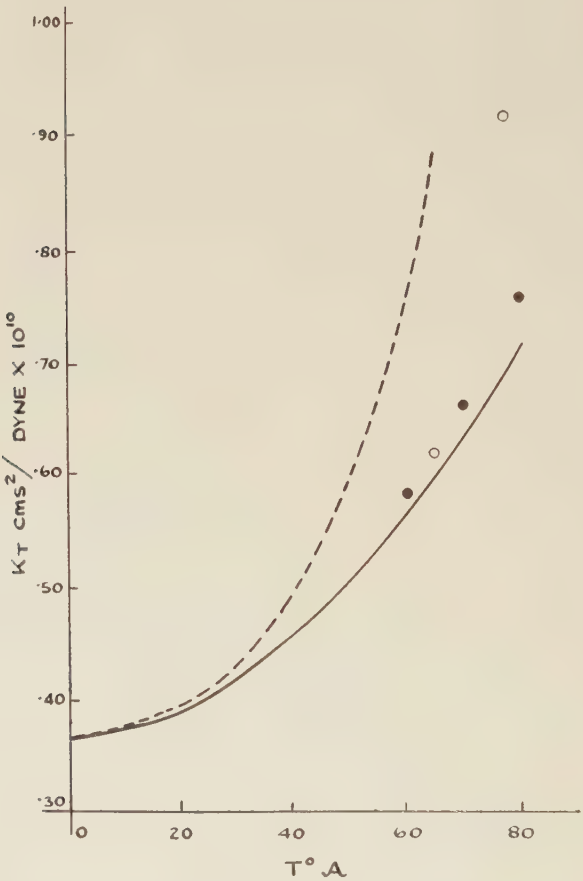
4.2. K_T *vs* T

At low temperatures there is little difference between K_T(H) and K_T(D) but large differences occur near the melting point where K_T(D) tends to infinity†. The few experimental results available favour K_T(H) rather

† Herzfeld and Goeppert-Mayer (1934) and Kane (1939) identify this with the melting point, but Frenkel (1946) points out that it only defines a certain condition of stability of a solid.

than K_T(D). Barker and Dobbs (1955) found K_T by measuring the velocity of ultrasonic waves in solid argon whilst Stewart (1955, 1956) found K_T using the piston displacement method. Although Stewart's

Fig. 2



Isothermal compressibility against temperature. ———— Henkel theory. — — — — — Debye theory. . ○ Experimental results—Stewart. ● Experimental results—Barker and Dobbs.

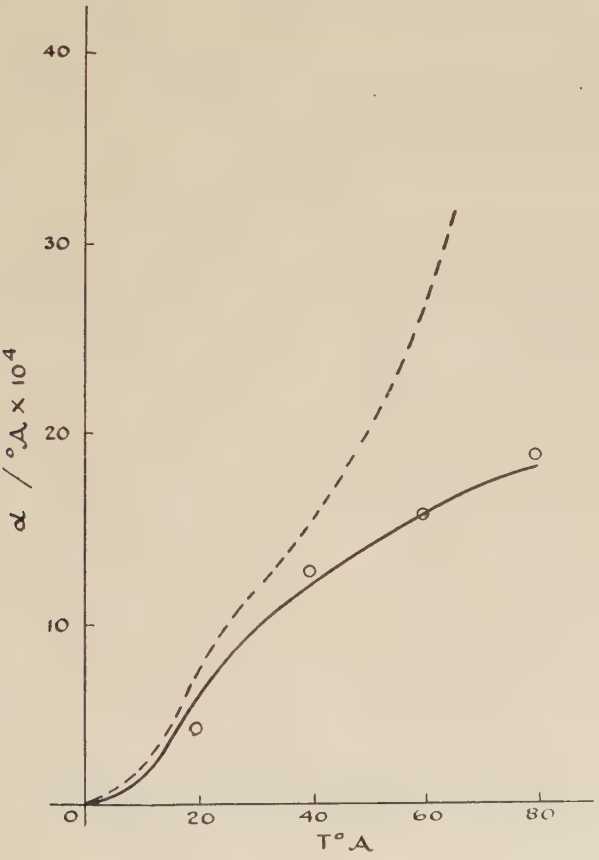
values at 65°K agree with Barker and Dobbs', there is a large discrepancy at 77°K but as Stewart himself points out his method of finding K_T is not trustworthy when K_T varies rapidly as it does at high temperature and low pressures. It should be pointed out that the values for K_T(H) given

here are slightly different from those originally communicated privately to Dobbs and Jones but this does not affect their conclusions.

4.3. α *vs* T

Again differences between α(D) and α(H) become appreciable at high temperatures and again α(H) agrees much better with experiment. The values of α(H) given here are also slightly different from those published by Dobbs and Jones but again make no difference in interpretation.

Fig. 3



Volume expansivity against temperature. ———— Henkel theory. — — — — — — Debye theory. ○ Experimental results—Dobbs *et al.*

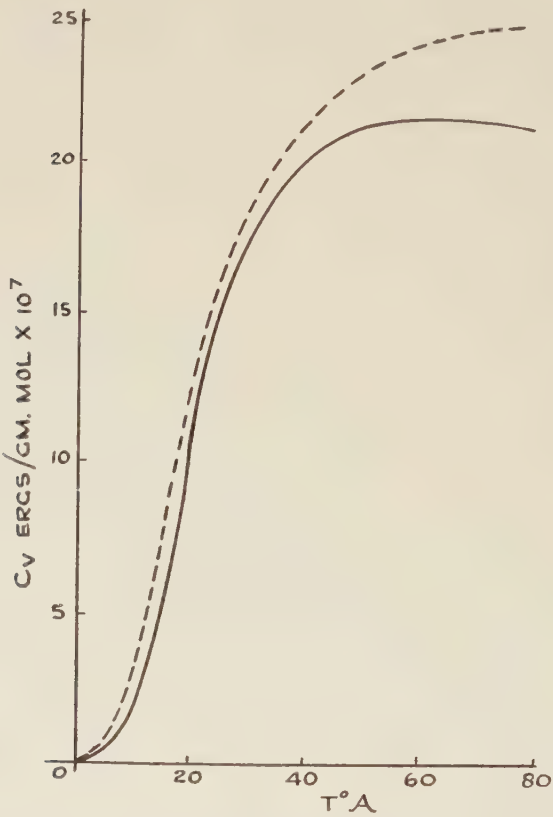
4.4. C_V and C_P *vs* T

At low temperatures C_V(D) is slightly greater than C_V(H). As the Henkel theory is just a modified Einstein theory this is just as might be

expected. But as the temperature rises C_V(H) does not approach $3R$ as C_V(D) and the Einstein specific heat do. This is undoubtedly due to the effect of anharmonicities which are expected to be more prominent at more elevated temperatures.

The C_P curves show that at low temperatures C_P(D) is in better agreement with experiment, whilst at high temperatures C_P(H) is better. This may be interpreted as follows. It is known that at low temperatures

Fig. 4

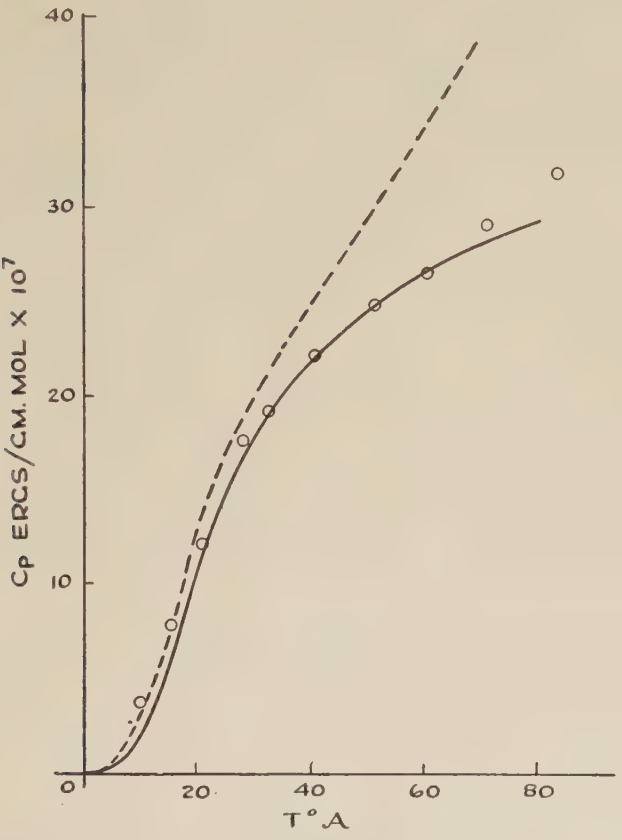


Specific heat at constant volume against temperature. ———— Henkel theory. — — — — — Debye theory.

the ordinary Einstein theory does not predict the behaviour of specific heat as well as does the Debye theory. Further as has already been seen anharmonic effects are small at low temperatures. Hence as the Henkel theory only modifies the Einstein theory by the inclusion of anharmonic effects, it might be expected that C_P(D) at low temperatures represents the experimental facts more closely. But as the temperature rises to

values greater than $\theta_D/5$ ($\simeq 16$ Å for argon) the differences between the Debye and the ordinary Einstein specific heats rapidly approach zero. Thus any difference between C_P(D) and C_P(H) at high temperatures must be attributed to anharmonicity, and here it is observed that the anharmonic theory represents C_P more closely.

Fig. 5

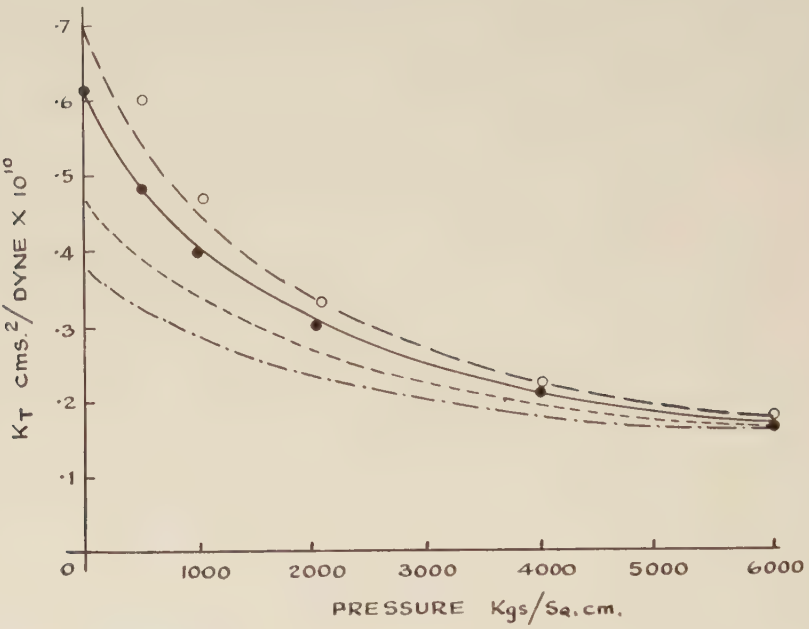


Specific heat at constant pressure against temperature. ———— Henkel theory. — — — — — Debye theory. ○ Experimental results—Clusius.

4.5. K_T *as a Function of Pressure*

It was found that at pressures greater than 6000 bars (1 bar = 0·98 atmospheres) that values of K_T at different temperatures became indistinguishable. The differences between K_T(H) and K_T(D) also rapidly approached zero, and the latter have not been illustrated here. The agreement with Stewart's results is good except at low pressures, and here the criticisms of § 4.2 apply.

Fig. 6



Isothermal compressibility as a function of temperature. — — — Henkel theory, $K_{T=77}$. ———— Henkel theory, $K_{T=65}$. – – – – – Henkel theory, $K_{T=40}$. —.—.—. Henkel theory, $K_{T=0}$. ○ Experimental results for $K_{T=77}$.—Stewart. ● Experimental results for $K_{T=65}$—Stewart.

Table 1 (*a*)

T Å	ρ(D) g/cm³	ρ(H) g/cm³	V(D) cm³	V(H) cm³	K_T(D) dynes⁻¹ cm²	K_T(H) dynes⁻¹ cm²
Omitted Factor	1	1	1	1	10^{-10}	10^{-10}
0	1·770	1·770	22·57	22·57	0·370	0·368
10	1·769	1·767	22·58	22·60	0·373	0·372
20	1·761	1·763	22·68	22·66	0·395	0·388
30	1·746	1·749	22·88	22·83	0·429	0·424
40	1·723	1·732	23·18	23·06	0·498	0·461
50	1·696	1·711	23·56	23·34	0·604	0·510
60	1·660	1·688	24·06	23·66	0·750	0·568
70	1·611	1·664	24·79	24·01	1·119	0·635
80	1·499	1·637	26·64	24·40	?	0·712

Table 1 (*b*)

T Å	C_V(D) joules mole	C_V(H) joules mole	C_P(D) joules mole	C_P(H) joules mole	α(D) Å $\times 10^{-4}$	α(H) Å $\times 10^{-4}$
0	0·0	0·0	0·0	0·0	0·0	0·0
10	3·44	1·54	3·46	1·57	1·77	0·73
20	12·72	10·97	13·27	11·32	6·91	5·48
30	18·35	16·82	20·23	18·15	10·83	9·06
40	21·05	19·96	25·51	22·48	14·39	11·23
50	22·61	20·53	29·49	24·43	18·78	13·07
60	23·48	21·28	34·82	26·56	24·28	14·53
70	24·03	21·27	39·37	28·31	37·35	16·31
80	24·03	20·97		29·33		17·36

Table 2

Molar volume (cm^3)	24·91	23·73	22·54	21·36	20·17	19·00
$T=0$						
P	−1790	−1090	77	1880	4800	9430
K(H)	0·907	0·565	0·367	0·243	0·162	0·108
$T=0$						
$T=20$ Å						
P	−1650	−980	160	1940	4840	9440
K(H)	0·955	0·584	0·372	0·245	0·163	0·108
$T=20$						
$T=40$ Å						
P	−1210	−540	571	2310	5160	9710
K(H)	0·966	0·596	0·380	0·249	0·165	0·110
$T=40$						
$T=65$ Å						
P	−630	80	1210	2960	5800	10330
K(H)	0·894	0·580	0·377	0·250	0·166	0·110
$T=65$						
$T=77$ Å						
P	−350	370	1530	3290	6130	10700
K(H)	0·850	0·566	0·373	0·248	0·166	0·110
$T=77$						

P, the pressure, is given in bars or dynes $\times 10^6$/cm^2.
K_T, the isothermal compressibility, is given in units of cm^2/dyne $\times 10^{-10}$.

§ 5. Conclusion

Various properties of solid argon have been evaluated employing two theories—the Debye theory and an anharmonic theory developed by Henkel from the Einstein model of a crystal. In all cases except for specific heats at low temperatures the anharmonic theory gives results agreeing more closely with available experimental data. This better agreement is most prominent at high temperatures. It would appear that anharmonic effects play a prominent part in determining the behaviour of solid argon, especially near the melting point. It is worth repeating the observations of Dobbs and Jones that more experimental data especially at low temperatures are required, and that an anharmonic theory based on a better model than that of Einstein is desirable.

Acknowledgments

The writer is grateful to Professor C. Domb of King's College, London, for many helpful discussions, and to Professor G. O. Jones and Dr. E. R. Dobbs of Queen Mary College, London, who communicated many results before publication. He is also indebted to D.S.I.R. for a maintenance grant for the period during which this work was done.

References

Barker, J. R., and Dobbs, E. R., 1955, *Phil. Mag.*, **46,** 1069.
Born, M., 1951, *Fest. Gott Akad. maths. Phys.* KL 1.
Born, M., and von Karman, T., 1912, *Z. Phys.*, **13,** 297; 1913, *Ibid.*, **14,** 15.
Clusius, K., 1936, *Z. phys. Chem.* B, **31,** 459.
Debye, P., 1912, *Ann. Phys.*, **39,** 789.
Dobbs, E. R., Figgins, D. F., Jones, G. O., Piercy, D. C., and Riley, D. P., 1956, Nature, *Lond.*, **178,** 483.
Dobbs, E. R., and Jones, G. O., 1957, *Rep. Prog. Phys.*, **20,** 516.
Domb, C., and Salter, L., 1952, *Phil. Mag.*, **43,** 1083.
Domb, C., and Zucker, I. J., 1956, *Nature, Lond.*, **178,** 484.
Einstein, A., 1907, *Ann. Phys.*, **22,** 180 ; 1911 a, *Ibid.*, **34,** 170 ; 1911 b, *Ibid.*, **35,** 679.
Frenkel, J., 1946, *Kinetic Theory of Liquids* (Oxford: University Press).
Henkel, J. H., 1955, *J. chem. Phys.*, **23,** 681.
Hooton, D. J., 1955 a, *Phil. Mag.*, **46,** 422; 1955 b, *Ibid.*, **46,** 433.
Herzfeld, K. F., and Goeppert-Mayer, M., 1934, *Phys. Rev.*, **46,** 995.
Kane, G., 1939, *J. chem. Phys.*, **7,** 603.
Lennard-Jones, J. E., and Ingham, A. E., 1925, *Proc. roy. Soc.* A, **107,** 146.
Stewart, J. W., 1955, *Phys. Rev.*, **97,** 578 ; 1956, *J. Phys. Chem. Solids*, **1,** 636.
Zucker, I. J., 1956, *J. chem. Phys.*, **25,** 915.

Some Magnetic Properties of Dilute Ferromagnetic Alloys II†

By B. W. Lothian, A. C. Robinson and W. Sucksmith
Department of Physics, University of Sheffield

[Received May 21, 1958]

Abstract

Previous experiments (Bate, G., Schofield, D. and Sucksmith, W. 1955) have been carried out on the magnetic properties of precipitates of dilute ferromagnetic alloys precipitated from solid solution in a non-ferromagnetic matrix. This process could be followed from the initial stages of superparamagnetism through single to multi-domain size of the aggregates, and the magnetic measurements corelated with particle growth. In the present communication, the work is extended to the production of precipitates with greater departure from spherical shape produced by cold drawing of suitable alloys of the ferromagnetics iron, nickel and cobalt. Magnetic measurements on the anisotropic specimens so produced are shown to give evidence for the distribution of particle shape, size and structure. The reverse magnetic field required to reduce the remanence to zero is shown to be an additional useful parameter in these determinations.

§ 1. Introduction

In their bulk polycrystalline forms, the ferromagnetic elements iron, nickel and cobalt are magnetically soft having coercive forces of no more than a few oersteds. However, if these elements are subdivided into sufficiently small particles they may exhibit coercive forces of several hundreds and even as high as thousands in certain cases. If such particles are sufficiently small to exist as single domains, changes in magnetization may only occur by the difficult process of rotation of the magnetization vector and not by the easy process of boundary displacement. The high coercive forces observed are due to the forces of anisotropy of the particles opposing this rotation of the magnetization vectors. The theoretical maximum values of coercivity associated with the three forms of anisotropy are shown in table 1 and are due to the work of Stoner and Wohlfarth (1948), Néel (1947 a) and Kittel (1949).

Numerous experimental investigations of the magnetic properties of finely divided ferromagnetic powders have been carried out. Bertaut (1953) and more recently Meikeljohn (1953) studied the dependence of coercive force on particle size and both obtained results supporting the theoretical work of Néel (1947 b). Considerable information regarding the magnetic properties of small particles has been obtained by studying a simple model consisting of small amounts of the ferromagnetic elements

† Communicated by the Authors.

dispersed in a non-magnetic matrix. Such a system may be realized by the precipitation of small ferromagnetic regions in alloys from super saturated solid solutions by annealing. Such solid solutions of a ferromagnetic and a non-ferromagnetic element should have a low solubility limit for the ferromagnetic element so that interaction between the ferromagnetic elements does not occur. Thus, a detailed investigation of the variation of the magnetic properties of copper rich copper–iron and copper–cobalt alloys with annealing after quenching was carried out by Bate *et al.* (1955). After quenching these alloys, which contained from 0 to 2% of the ferromagnetic constituent, were not ferromagnetic. However, during annealing at temperatures ranging from 350°C to 500°C for the copper–cobalt alloys and from 500°C to 800°C for the copper–iron alloys, these alloys became weakly ferromagnetic developing both remanence and coercivity. With continued annealing, the coercivities of these alloys increased to maximum values of 250 oe and 400 oe for the copper–cobalt and copper–iron alloys respectively, at which stage the largest proportion of regions behaving as single domains was present. The formation of multidomain particles by further precipitation was accompanied by a decrease in the coercivity of the alloys.

Table 1

	Crystal		Shape		Strain	
	random	oriented	random	oriented	random	oriented
Iron	160	500	5100	10700	300	600
Nickel	60	185	1450	3000	2000	4000
Cobalt	2000	6000	4300	8900	300	600

All values are shown in Oersteds.

Although the maximum coercive forces observed for the copper–cobalt and copper–iron alloys were only 250 oe and 400 oe respectively, it was found that the reverse field required to reduce the remanence to zero was usually greater than 1000 oe and often as high as 1500 oe for both alloys. In ordinary ferromagnetic materials, this reverse field (H_r) is usually not more than about 30% more than the coercive force (H_c). A high value of the ratio may be taken to indicate that the hysteresis loop is the resultant of high and low coercivity contributions. Furthermore, the ratio of the remanent magnetization to the saturation value provides information concerning the distribution of the magnetization vectors amongst the contributing particles.

The work described in the present paper dealt with three dilute ferromagnetic alloys consisting of iron in copper, nickel in gold and cobalt in copper. In order to modify the shape and orientation of the domains, suitable specimens of the alloys were cold drawn, the magnetic properties being investigated at various stages of reduction.

§ 2. Experimental Techniques

2.1. *Preparation of the Alloys*

The alloys were prepared in 100 gram melts by melting the 99·9% pure constituents in a high frequency induction furnace in an atmosphere of argon at a pressure of 5 cm of mercury. Alumina crucibles were used in the preparation of the copper–iron and gold–nickel alloys, whilst the crucibles used for the copper–cobalt alloys were made from pure Acheson graphite.

2.2. *Heat Treatment*

The copper–iron and copper–cobalt alloys were forged at bright red heat to bars approximately 1 cm in diameter and 7 cm in length. Forging at bright red heat of the gold–nickel alloys so prepared was found to produce large cracks. Consequently, the alternative procedure of melting these alloys in a mould of the required shape in the hearth of an argon arc furnace was adopted. The bars obtained by these processes were then heat treated at 1000°c before quenching. Initially, all the alloys were quenched in water, but in the case of the copper–cobalt alloys this was found to produce surface strains which affected the rate of precipitation. This was remedied in subsequent work by oil quenching.

Investigations of the effect of further heat treatment on the magnetic properties of these alloys were carried out on specimens in the form of discs 5 mm in diameter and 0·5 mm thick, these having been annealed and quenched to remove surface strain produced in preparation. The appropriate heat treatments were carried out with the specimens sealed into small evacuated hard glass tubes.

After the bar had been heat treated, it was turned down to 8 mm diameter and threaded at one end in preparation for cold drawing. Held by means of a threaded steel tube and steel rod, the bar was cold drawn through dies. By this method the diameter of the bar was reduced in steps of $\frac{1}{4}$ mm to 4 mm and then in $\frac{1}{8}$ mm steps down to 2 mm. At various stages of reduction, specimens were cut from the bar with the plane of the disc parallel to the direction of drawing. A small arrow was scratched on the surface of each disc to coincide with the direction of drawing.

2.3. *Magnetic Measurements*

2.3.1. *Determination of the hysteresis loops*

The demagnetization curves were determined by measuring the tractive force on a magnetized specimen produced by means of a field gradient. This was achieved by means of a torsion balance described by Bate *et al.* (1955).

The force F exerted on a specimen of mass M and intensity of magnetization σ in a field gradient dH/dZ is given by $\sigma M(dH/dZ)$.

$$F = \sigma M \frac{dH}{dZ} .$$

The coils magnetizing the specimen to its intensity σ and these producing the necessary field gradient are separately controlled and collinear.

This force is balanced by the couple exerted by the torsion fibre of constant c, and thus we obtain

$$c\theta = \sigma M \frac{dH}{dZ} d$$

where θ is the angle of rotation of the torsion head necessary to restore the balance arm of length d to its zero position.

The arrangement of the apparatus and experimental technique were similar to those described by Bate *et al.*

2.4. *Saturation Measurements*

The intensity of magnetization of specimens in fields of up to 18 000 oe were determined using a Sucksmith magnetic ring balance (Sucksmith 1929).

§ 3. Experimental Results

3.1. *Copper–Iron Alloys*

The investigations of the effects of heat treatment and of cold work after heat treatment on the magnetic properties of copper rich copper–iron alloys were carried out on alloys containing 0·5%, 1·0% and 2·0% by weight of iron. After annealing at 1000°c for several hours followed by quenching, the 0·5% and 1·0% alloys were not ferromagnetic but the 2·0% alloy showed slight ferromagnetism, having a remanence of $\sigma = 0{\cdot}015$ (e.m.u. per gram). If all the iron was dissolved by the heat treatment, as will be expected from the equilibrium diagram, some was precipitated as body centred iron on quenching.

The effect of further heat treatment at 650°c on the magnetic properties of these alloys was then studied and the results are summarized in table 2.

Table 2. The Effect of Heat Treatment at 650°c on the Coercivity and Field Required to reduce Remanence to Zero

Annealing time in hours	H_c (oe)			H_r (oe)		
	2	5	12	2	5	12
½% Fe–Cu	320	390	160	750	450	530
1% Fe–Cu	420	480	290	800	820	650
2% Fe–Cu	180	210	100	660	830	1250

As seen in table 2 the initial coercive force of these alloys in general increased to a maximum after about 5 hours heat treatment and then decreased with further heat treatment. Similar variations of the field required to reduce remanence to zero were observed but the period of

heat treatment required to give the maximum values was found to be dependent on the composition of the alloy. Although the maximum values of H_C and H_R were almost identical with those observed by Bate *et al.* (1955) the periods required to achieve these maxima were much shorter in the present work, e.g. the maximum value for the coercivity of a 1% alloy was observed after 65 hours at 650°C compared with a period of only 5 hours in the present work. This difference in the rates of precipitation is due to a difference in the preparation of the two sets of alloys. The earlier alloys were quenched directly after forging whereas the present alloys were heat treated for a further 24 hours at 1000°C before quenching. These differences can therefore be attributed to the stresses set up in forging.

The effect of cold drawing on the magnetic properties of alloys which had been heat treated for periods of 2, 5 and 12 hours at 650°C after quenching was then studied and a summary of the observed effects is given in table 3. Cold drawing of quenched alloys produced only small changes in the magnetic properties, particularly in so far as no anisotropy developed.

The results of magnetic measurements on 0·5%, 1% and 2% Fe–Cu alloys are shown in tables 3 (*a*), (*b*) and (*c*). After heat treatments of 2, 5 or 12 hours at 650°C it was found that even small reductions in cross-sectional area, e.g. 20%, produced substantial increases in both the remanence and the saturation intensity of these alloys. As seen in the appropriate columns of table 3, the percentage increase in both σ_{RP} and σ_S was greatest for the initial reductions and tended to a maximum after about 75%. This shows that considerable increase in the ferromagnetism has occurred, due to a transformation into the ferromagnetic α form, by cold work, of the γ iron originally precipitated by heat treatment. In addition to this further precipitation, cold drawing of these alloys after heat treatment resulted in a highly anisotropic state with the coercive force parallel to the direction of drawing being very much greater than in the perpendicular direction. As can be seen from the values of H_{CP} and H_{CN}, although H_{CP} approaches the value of the field required to reduce remanence to zero, H_{RP}, as the reduction is increased, very little effect on the values of H_{CN} was observed.

Although H_{CN} was hardly affected by continued reduction the effect on H_{RN} was very similar to that on the value of H_{RP}. However, the magnitude of the effect of cold drawing on Hr was much smaller than on H_{CP}, suggesting that cold drawing only has a small effect on the initial ferromagnetic precipitate and that the high coercive forces are due to orientation of the new ferromagnetic material.

The ratio σ_R/σ_S, an indication of the amount of reversible magnetization which may occur, is also shown in table 3. Before drawing, the value of this ratio was low, e.g. $\sim 0{\cdot}1$–$0{\cdot}4$, indicating a large reversible component. Drawing of the alloys was found in general to cause a progressive decrease in the ratio σ_{RN}/σ_S, whereas the ratio σ_{RP}/σ_S increased, until, with 90% reduction, a value approaching unity was produced. These results support the evidence from the values of H_C and H_R that rotation of the precipitate without much change in shape is taking place.

Table 3

(*a*) The Variation of Magnetic Properties of a 0·5% Iron–Copper Alloy with Cold Drawing after 5 hours Heat Treatment at 650°c

Diam. mm.	% Red	H_{CP}	H_{RP}	$\frac{H_{RP}}{H_{CP}}$	H_{CN}	H_{RN}	$\frac{H_{RN}}{H_{CN}}$	σ_{RP}	σ_{RN}	σ_S	$\frac{\sigma_{RP}}{\sigma_S}$	$\frac{\sigma_{RN}}{\sigma_S}$
8	0	390	450	1·2	390	450	1·2	0·0045	0·0045	0·0112	0·40	0·40
7	24	560	660	1·2	285	660	2·3	0·174	0·0483	0·230	0·76	0·21
6	44	760	860	1·1	320	889	2·8	0·270	0·0607	0·303	0·89	0·20
5	61	875	940	1·1	360	1150	3·2	0·309	0·0652	0·332	0·93	0·195
4	75	1000	1075	1·1	320	1280	4·0	0·326	0·0584	0·362	0·91	0·16
3	86	920	1020	1·1	310	1200	3·9	0·292	0·0562	0·349	0·84	0·16
1·87	94	1100	1260	1·1	310	1350	4·4	0·315	0·0529	0·360	0·88	0·15

(*b*) The Variation of Magnetic Properties of a 1% Iron Copper Alloy with Cold Drawing after 5 hours Heat Treatment at 650°c

Diam. mm.	% Red	H_{CP}	H_{RP}	$\frac{H_{RP}}{H_{CP}}$	H_{CN}	H_{RN}	$\frac{H_{RN}}{H_{CN}}$	σ_{RP}	σ_{RN}	σ_S	$\frac{\sigma_{RP}}{\sigma_S}$	$\frac{\sigma_{RN}}{\sigma_S}$
8	0	486	820	1·7	486	820	1·7	0·0495	0·0495	0·112	0·44	0·44
7	24	480	630	1·3	280	640	2·3	0·272	0·102	0·462	0·59	0·22
6	44	850	1000	1·2	550	1150	2·1	0·529	0·197	0·775	0·68	0·25
5	61	950	1250	1·3	460	1150	2·5	0·705	0·179	0·933	0·76	0·19
4	75	1110	1300	1·2	430	1200	2·8	0·820	0·146	1·04	0·79	0·14
3	86	1265	1380	1·1	410	1390	3·4	0·910	0·152	1·06	0·86	0·14
2·12	94	1170	1260	1·1	300	1160	3·9	0·973	0·125	1·06	0·92	0·12

(*c*) The Variation of Magnetic Properties of a 2% Copper-Iron Alloy with Cold Drawing after 5 hours Heat Treatment

Diam. mm.	% Red	H_{CP}	H_{RP}	$\frac{H_{RP}}{H_{CP}}$	H_{CN}	H_{RN}	$\frac{H_{RN}}{H_{CN}}$	σ_{RP}	σ_{RN}	σ_S	$\frac{\sigma_{RP}}{\sigma_S}$	$\frac{\sigma_{RN}}{\sigma_S}$
8	0	210	830	4·5	210	830	4·5	0·063	0·063	1·26	0·048	0·048
7	24	520	670	1·3	250	740	3·0	0·905	0·307	2·88	0·314	0·120
6	44	740	910	1·2	306	1200	3·9	1·47	0·388	3·24	0·455	0·120
5	61	700	860	1·2	200	950	4·7	1·69	0·281	3·57	0·474	0·079
4	75	790	1000	1·3	365	1300	3·6	1·73	0·472	3·63	0·477	0·130
3	86	840	1100	1·3	310	1210	3·9	1·54	0·517	3·63	0·477	0·142
1·87	94	800	1070	1·3	310	1320	4·3	1·69	0·506	3·51	0·481	0·144

H_{CP}—Coercivity parallel to the direction of drawing.
H_{CN}—Coercivity perpendicular to the direction of drawing.
H_{RP}—Reverse field required to reduce remanence to zero parallel to drawing.
H_{RN}—Reverse field required to reduce remanence to zero perpendicular to drawing.
σ_R—Remanence.
σ_S—Saturation intensity.

In general the contributions to the coercive force arising from all three forms of anisotropy will be influenced by cold drawing. Crystal anisotropy will be affected according to the structures of the two phases concerned. Thus in the case of face-centred cubic structures the ⟨100⟩ and ⟨111⟩ axes, whereas in body centred cubic materials the ⟨110⟩ axes are aligned parallel to the direction of drawing, the contributions from shape anisotropy may be two-fold. There may be elongation of the particles caused by the drawing process, and there may be orientation of the particles without deformation. From the measurements of coercivity and magnetization referred to above, it appears unlikely that there is much change of shape, but the increase of coercive force in the direction of drawing (as well as the decrease in the perpendicular direction) would both result from the cold work aligning the long axes of the particles along the direction of drawing. Although strain effects will be present, it does not seem likely that will play a major part in so far as directional properties are concerned. Of the possible causes of the changes in magnetic properties, the evidence points towards the following explanation. Before cold work, there is a considerable amount of face-centred iron which is coherent with the copper rich matrix. During the drawing process, these copper rich grains, and probably those coherent with them, are turned into the direction of drawing and at the same time transformed into body-centred cubic iron, so that finally we have ferromagnetic grains most of which have been elongated to a greater extent than prior to drawing. At the same time, most of these have their long axes along the direction of drawing.

3.2. *Copper–Cobalt Alloys*

The variation with heat treatment of the magnetic properties of quenched copper-rich copper–cobalt alloys containing 1% and 1½% by weight of cobalt was investigated. The effect of cold drawing after heat treatment was also studied. Some measurements were carried out on a 4% cobalt–copper alloy.

By quenching the 1% and 1½% alloys in oil after 24 hours annealing at 1000°C all the cobalt was retained in solid solution and the alloys exhibited paramagnetic behaviour with no signs of ferromagnetism. After annealing at temperatures ranging from 350°C to 450°C, the ferromagnetic characteristics of remanence and coercivity were developed. The variations of magnetic properties for a 1% alloy are summarized in table 4. The 1½% alloy behaved in a very similar manner.

Short periods of heat treatment were found to develop low coercive forces of about 50 oe which increased to a maximum on the continued heat treatment. This maximum was dependent on the temperature of heat treatment, e.g. for the 1% alloy the maximum coercive force for heat treatment at 300°C was 145 oe, whereas at 380°C it was 65 oe. Although the values of H_C were comparatively low, i.e. compared with the results for iron, the values of H_R were very high with values of up to 1300 oe being observed. As seen in table 4 the values of the ratio H_R/H_C were very high with values always greater than seven, and often greater than 10.

This suggests that single domain behaviour is involved in the high values of H_R, and also that a considerable number of the precipitated regions were magnetically soft, having coercivities less than 100 oe. Particles the size of which is either less than, i.e. subdomain, or much greater than, i.e. multidomain, single domain size constitute 'soft' regions. Such regions will also be associated with the very low values of σ_R/σ_S which indicates a large proportion of reversible magnetization.

Table 4. The Variation of Magnetic Properties with Heat Treatment at 350°C of a 1% Copper–Cobalt Alloy

Time (hr)	H_C	H_R	$\frac{H_R}{H_C}$	σ_R	σ_S	$\frac{\sigma_R}{\sigma_S}$
4	90	920	10	0·0009	0·0250	0·036
8	103	920	8·9	0·0014	0·0275	0·051
22	112	890	8·0	0·0023	0·0400	0·058
45	125	900	7·2	0·0031	0·0450	0·061
93	112	890	7·9	0·0036	0·0700	0·052

Although the general variation of the magnetic properties of these alloys with heat treatment was similar to that observed by Bate *et al.* (1955), the rate of precipitation and the maximum coercivities were different in the two cases. This difference may be accounted for by the fact that, in the earlier work, heat treatments were carried out on specimens immediately after cutting from the forged bar, whereas in the present work specimens were quenched after preparation to remove surface strain before the heat treatments.

The results of cold drawing on a $1\frac{1}{2}$% cobalt–copper alloy are shown in table 5. Before drawing, the specimens were magnetically isotropic with low values of H_C and high values of H_R. The drawing produces virtually no change in either the coercivity, nor the value of the field required to reduce the remanence to zero, irrespective of the direction concerned. The values of the saturation intensity show that no further precipitation occurred with successive cold work. Little magnetic anisotropy is developed in this system, though there is certainly indication of greater irreversibility in the drawing direction. Microscopic examination of specimens which had been polished and etched showed that cold drawing did not affect the shape of the smaller particles, whilst the larger particles which were elongated in the direction of drawing. Results for a 1% Co–Cu follow closely behaviour similar to the $1\frac{1}{2}$% alloy.

The results for the 4% copper–cobalt alloys were however somewhat different from those for the 1% and $1\frac{1}{2}$% alloys in that after quenching in oil from 1000°C this alloy was ferromagnetic with a saturation magnetization of $\sigma_S \sim 5$ e.m.u./g. This is not unexpected since the solubility limit only reaches 4% at about 1000°C. The effects of cold drawing on the

quenched alloy were investigated and the results observed are summarized in table 6.

Cold drawing of the 4% alloy increased the low value of H_C observed for the undrawn material by similar amounts in both the parallel and perpendicular directions for reductions of up to 80%. However, on further reduction in area H_{CP} was increased by a greater amount than

Table 5. The Effect of Cold Drawing on the Magnetic Properties of a $1\frac{1}{2}$% Cobalt–Copper Alloy after 46 hours Heat Treatment at 400°C

Diam. mm.	% Red	H_{CP}	H_{RP}	$\frac{H_{RP}}{H_{CP}}$	H_{CN}	H_{RN}	$\frac{H_{RN}}{H_{CN}}$	σ_{RP}	σ_{RN}	σ_S	$\frac{\sigma_{RP}}{\sigma_S}$	$\frac{\sigma_{RN}}{\sigma_S}$
8	0	26	700	27	26	700	27	0·014	0·014	0·450	0·031	0·031
7	24	18	675	37	18	790	44	0·007	0·0077	0·320	0·022	0·024
6	44	17	800	47	14	670	47	0·0038	0·0031	0·240	0·014	0·013
5	61	21	500	24	25	890	35	0·0066	0·0029	0·225	0·029	0·013
4	75	17	700	41	18	725	40	0·0072	0·0031	0·260	0·028	0·012
3	86	23	700	30	25	730	29	0·0080	0·0028	0·250	0·032	0·011

Table 6. The Effect of Cold Drawing on the Magnetic Properties of a 4% Copper–Cobalt Alloy after Quenching in Oil

Diam. mm.	% Red	H_{CP}	H_{RP}	$\frac{H_{RP}}{H_{CP}}$	H_{CN}	H_{RN}	$\frac{H_{RN}}{H_{CN}}$	σ_{RP}	σ_{RN}	σ_S	$\frac{\sigma_{RP}}{\sigma_S}$	$\frac{\sigma_{RN}}{\sigma_S}$
8	0	16	220	14	16	220	14	0·16	0·16	5·30	0·030	0·030
7	24	33	565	17	27	615	23	0·50	0·18	5·20	0·096	0·035
6	44	54	935	18	48	1050	22	0·77	0·21	5·20	0·150	0·040
5	61	77	1300	17	66	1450	22	0·82	0·21	5·10	0·160	0·041
4	75	94	1470	16	83	1650	20	0·85	0·21	5·00	0·170	0·042
3	86	80	1520	19	80	1750	22	0·73	0·16	4·80	0·150	0·034
1·87	96	310	1000	3·4	150	1225	8·0	1·30	0·24	4·30	0·300	0·055

H_{CN}, the actual values of H_{CP} and H_{CN} after 96% reduction were 310 oe and 150 oe respectively. The drawn material was almost isotropic with respect to the values of H_R and apart from the values observed for material which had received 96% reduction in area, the values of H_R/H_C remained almost constant. Although no appreciable effect on the saturation intensity nor the values of σ_{RN} was observed, the value of σ_{RP} was increased by a factor of 10 by 96% reduction. This gave a corresponding increase in the value of σ_{RP}/σ_S but even after 96% reduction the value was only 0·30.

For such small amounts of cobalt as were involved in the alloys studied here it is not possible to obtain an x-ray determination of the structure of the precipitated regions. Although the evidence is in favour of the structure of the precipitated regions being face-centred cubic, the possibility of it being hexagonal does exist. The two most interesting features of the magnetic properties of the alloys were the high values of H_R/H_C and the low values of σ_R/σ_S. Considering the former, while any one of the three forms of anisotropy may give contributions sufficiently high to explain the low H_C values, it is not easy to explain the values of H_R. However, these high values may be explained by assuming that a small amount of the precipitate consists of single domain hexagonal cobalt regions from which quite considerable contributions to the H_C values may arise from crystal anisotropy. Since the hexagonal basal plane tends to lie along the drawing direction for many hexagonal metals, this will tend to enhance the contribution to the coercivity from this source. Furthermore, as the coercive force of a random assembly of single domain cobalt particles having an axial ratio of only 1·2 may be as high as 600 oe, a certain contribution is almost inevitably due to shape. If the remainder of the precipitate is magnetically soft, this might account for the low values of σ_R/σ_S involved in determining the values of H_R. The fact that the remanence along the direction of drawing is appreciably higher than the corresponding value in a perpendicular direction points to change in the shape factor caused by drawing. Owing to the low disregistry of the two precipitates, it would appear likely that elongation of the particles might be a more important factor than for the case of the copper iron system.

3.3. *Gold–Nickel Alloys*

Investigations of the variation of the magnetic properties of gold-rich gold–nickel alloys with heat treatment and also with subsequent cold drawing were carried out on alloys containing 6% and 9% by weight of nickel.

In early experiments, heat treatments were carried out on specimens immediately after cutting from the quenched forged bar. Before heat treatment these specimens were not ferromagnetic, but became so during heat treatment. Later, as for copper-cobalt alloys, all specimens were quenched from 1000°C in water before heat treatment and all subsequent heat treatments were carried out on such quenched specimens.

The effect of heat treatment at temperatures ranging from 300°C to 400°C was studied for both alloys and the results obtained by heat treatment at 350°C of the 9% alloy are summarized in table 7.

After short periods of heat treatment, the materials which were then magnetically isotropic, served coercive forces of about 200 oe. With further heat treatment this value increased to a maximum of about 300 oe and then decreased. The values of H_R showed similar variations to those of H_C and although H_R was greater than H_C by 100–150 oe, the difference

between the two was not as large as observed for other dilute ferromagnetic alloys. The value of H_R/H_C was only slightly affected by heat treatment, decreasing from 1·9 down to 1·2 and then remaining constant. For the 6% alloy H_R/H_C was constant having a value of 1·3.

Table 7. The Variation of Magnetic Properties with Time of Heat Treatment at 350°C of a 9% Gold–Nickel Alloy

Time (hr)	H_C	H_R	$\frac{H_R}{H_C}$	σ_R	σ_S	$\frac{\sigma_R}{\sigma_S}$
2	210	400	1·90	0·0004	0·010	0·040
4	235	410	1·75	0·0010	0·012	0·083
8	275	410	1·50	0·0025	0·014	0·180
16	285	385	1·30	0·0263	0·068	0·390
32	295	360	1·20	0·1180	0·285	0·410
64	230	280	1·20	1·350	2·900	0·470

Table 8. The Effect of Cold Drawing on the Magnetic Properties of a 9% Gold–Nickel Alloy after 30 hours Heat Treatment at 350°C

Diam. mm.	% Red	H_{CP}	H_{RP}	$\frac{H_{RP}}{H_{CP}}$	H_{CN}	H_{RN}	$\frac{H_{RN}}{H_{CN}}$	σ_{RP}	σ_{RN}	σ_S	$\frac{\sigma_{RP}}{\sigma_S}$	$\frac{\sigma_{RN}}{\sigma_S}$
8	0	305	410	1·3	305	410	1·3	0·186	0·186	0·440	0·42	0·42
7	24	490	890	1·8	645	735	1·1	0·121	0·279	0·520	0·23	0·54
6·5	34	450	800	1·8	605	690	1·1	0·108	0·208	0·400	0·27	0·52
6·0	44	430	685	1·6	570	675	1·2	0·156	0·256	0·490	0·32	0·52
5·5	53	385	585	1·5	510	605	1·2	0·212	0·226	0·550	0·39	0·49
4·5	69	365	570	1·5	515	630	1·2	0·182	0·307	0·610	0·30	0·50
4·0	75	305	444	1·4	410	540	1·3	0·242	0·236	0·570	0·42	0·41
3·0	86	295	485	1·6	450	575	1·3	0·142	0·202	0·450	0·32	0·45
2·1	93	300	400	1·3	350	500	1·4	0·257	0·190	0·530	0·48	0·36

Although microscopic examination of specimens after polishing and etching showed the particles to be spherical with sizes varying from well below to well above the critical single domain size, very small deviation from spherical shape is sufficient to give coercivities ~ 100 oe. Due to this low value, contributions to the coercivity of these alloys may arise from any of the different sources of anisotropy. The value of 1·3 for the ratio of H_R/H_C is not significant, since very small amounts of the low coercivity regions are required to give this value for H_R/H_C (Miekeljohn 1953).

As seen in table 7 the values of σ_R/σ_S increase from a very low value after short periods of heat treatment to values approaching 0·5 after

further heat treatment. This is in agreement with the value of $\sigma_R/\sigma_S = 0{\cdot}5$ predicted by Stoner and Wohlfarth (1948) for a random assembly of single domain particles.

The results for the 6% alloy were very similar to those for the 9% alloy. The maximum coercivities were obtained at temperatures of 350°C and 380°C for the 9% and 6% alloys respectively.

After the alloys had been heat treated to give the maximum coercive force, the effect of cold drawing on their magnetic properties was studied. The results obtained are summarized in table 8.

Before cold drawing, these alloys were isotropic with values of H_C and H_R of ~ 300 oe and 400 oe respectively. Cold drawing of the alloys resulted in anisotropic material with increased values of H_C and H_R both parallel and perpendicular to the direction of drawing. The values of H_C and H_R both increased to maximum after 25% reduction and on further reduction decreased. As seen in the respective columns of table 8, H_{CN} was always greater than the corresponding value of H_{CP}, the difference at their peak values being about 150 oe. The drawing was found initially to increase the value of H_{RP}/H_{CP} whilst that of H_{RN}/H_{CN} was decreased, these values remaining reasonably constant during the reduction. As has been pointed out above, both the ⟨100⟩ and ⟨111⟩ directions tend to align along the direction of drawing for face-centred cubic material, the proportions following one particular direction varying from one metal to another to that given above. The effect of drawing will therefore tend to change from a random distribution. If the proportions are of the same order of magnitude then the new distribution will show no particular preponderance for any crystal axis to lie either parallel or perpendicular to the drawing direction, i.e. approximate isotropy.

The results for the 6% alloy were very similar to those observed for the 9% alloy.

As the preferred orientation assumed by materials during cold drawing is a gradual process which may not be completed until the material has been heavily reduced in area. it appears that the initial large increases in the values of H_C and H_R are not associated with this type of process. Considering the values of coercive force associated with the various forms of anisotropy as shown in table 1, it appears that strain may play an important role in these initial increases. After a certain amount of reduction, further drawing will destroy any coherency between the precipitate and matrix. This process will probably be accompanied by a reduction in strain, and hence coercive force. The observed reduction in the values of H_C and H_R with reductions greater than 25% may be associated with the destruction of coherency and the orientation of the precipitate according to the double drawing texture of face-centred cubic material. This view is supported by the facts that the decrease was gradual and the materials became gradually more isotropic as the reduction in area progressed.

§ 4. Discussion

The variation of the magnetic properties with heat treatment was very similar for the three alloy systems studied and may be explained in terms of the growth of small ferromagnetic regions in a non magnetic matrix. The maximum coercive forces observed for these three alloy systems suggest that the higher coercive forces are associated with large disregistry. The values of $H_{C\max}$ and the corresponding disregistry values are given in table 9.

Table 9

	$H_{C\max}$	Disregistry
Fe–Cu	480 oe	25%
Co–Cu	150 oe	1·5%
Mi–Au	325 oe	10%

As discussed by Geisler (1953) large disregistry leads to a tetragonal structure as opposed to a simple cubic structure for coherent growth of the precipitate. In these alloys this gives rise to large internal strains and larger values of the crystal anisotropy constants than observed for simple cubic material. This means larger contributions to the coercive force from the relative anisotropies as shown in table 1.

Cold drawing of these alloys after heat treatment resulted in materials which were magnetically anisotropic, the anisotropy being more pronounced in the copper–iron alloys, where $H_{C\mathrm{P}} \gg H_{C\mathrm{N}}$ than in the other two systems studied. It is interesting to note that for materials involving precipitated regions of face-centred cubic structure the anisotropy produced by cold drawing was much less than for the one involving regions of body-centred cubic structure. Considering this and the fact that apart from the 4% copper–cobalt alloy little preferred shape orientation, or shape deformation was observed, it may be that the anisotropy produced was associated with crystal anisotropy and the drawing texture of the materials involved. Only in the case of the copper–iron alloys did cold drawing cause further precipitation of the ferromagnetic region. This was due to the transformation of the non-magnetic γ iron to the ferromagnetic α iron which is body-centred cubic.

Preliminary microscopic examinations were carried out on all three alloy systems. For the iron–copper alloys, the electron microscope photographs did not indicate any marked anisotropy, but the average diameter of the particles at maximum coercivity was about ten times larger than the generally accepted value for isolated single domain particles of iron. For the cobalt–copper alloys it appeared that cold drawing had little effect on the shape of the smaller particles, but larger ones—again considerably bigger than simple domain size—were elongated in the direction of drawing. In the case of the nickel–gold system, microscopic examination proved more difficult on account of the extreme softness of the matrix, but optical

microscopy revealed that as in the case of cobalt–copper alloys the larger particles were elongated in the drawing direction. The most important generalization to be drawn from these preliminary experiments indicated that the average particle size is much higher than would have been expected from single domain considerations. This and other aspects of the results are being further investigated.

Although it is true to say that contributions to the coercive force of these alloys, both drawn and undrawn, will arise from each of the three forms of anisotropy, it is not yet possible to make a satisfactory estimation of the relative contributions from each. However, further advance towards such an estimation may be made from further work, such as study of the temperature dependence of the magnetic properties of these alloys.

Acknowledgments

We are grateful to Dr. E. O. Hall and Professor R. W. K. Honeycombe for their help with the preparation and examination of specimens by means of the optical and electron microscopes.

References

Bate, G., Schofield, D., and Sucksmith, W., 1955, *Phil. Mag.*, **46,** 621.
Bertaut, F., 1953, *C. R. Acad. Sci., Paris*, **229,** 417.
Geisler, A. H., 1953, *Rev. mod. Phys.*, **25,** 316.
Kittel, C., 1949, *Rev. mod. Phys.*, **21,** 541.
Meikeljohn, W. H., 1953, *Rev. mod. Phys.*, **25,** 302.
Néel, L., 1947 a, *C.R. Acad. Sci., Paris*, **224,** 1488; 1947 b, *Ibid.*, **224,** 1550.
Stoner, E. C., and Wohlfarth, E. P., 1948, *Phil. Trans. roy. Soc.* A., **240,** 559.
Sucksmith, W., 1929, *Phil. Mag.*, **8,** 158.

A Note on Transition Metal Alloys†

By C. W. Haworth and W. Hume-Rothery
Department of Metallurgy, University Museum, Parks Road, Oxford

[Received May 14, 1958]

Abstract

The composition limits of the different crystal structures found in the alloys of Transition Elements of Groups VA to VIIIC are discussed. The alloys of these elements give rise to phases with crystal structures of the σ, c.p. hexagonal, αMn, and βW types. The composition limits of these are often related to the Group Numbers of the constituent elements. If the equilibrium diagrams are drawn with the element of lower Group Number on the left, passage across the diagram from left to right results in the occurrence of different phases in the following order:

b.c. cube→βW→σ→αMn→c.p. hex.→f.c. cube.

In all systems one or more of the above phases is absent, but the characteristic order is retained. In the alloys of elements of Group VIIIC with the elements in the middle of the transition series, there is a general tendency for the primary solid solution in the element of Group VIIIC to be greater than that in the element of the earlier Group, and an explanation of this is advanced.

§ 1. Types of Structure

The object of the present note is to discuss the composition limits of the different crystal structures found in the alloys of the Transition Elements of Groups VA–VIIIC. The electron theory of these metals is still a matter of acute controversy, and only a few of the binary equilibrium diagrams are accurately known. However, when the existing diagrams are compared systematically, some general principles or tendencies are now apparent.

For convenience we shall describe these first in terms of the Group Number of the elements, and regard this as increasing from IV (Ti, Zr, Hf) to VIIIC (Ni, Pd, Pt). All diagrams are drawn with the Group Number increasing from left to right. In these diagrams, on passing from left to right, the number of electrons outside the inert gas shell increases, and the number of holes in the d shell decreases. If a viewpoint of the Pauling type is adopted, the number of vacancies in the atomic orbitals decreases on passing from Group VI to Group VIIIC; there are no atomic orbitals in the elements of Groups IV or V.

In the Second and Third Long Periods the crystal structures of the pure metals show a clear change with increasing Group Number from

b.c. cube (Groups V, VI)→c.p. hex. (Groups VII, VIIIA)→
f.c. cube (Groups VIIIB and C).

† Communicated by the Authors.

In the First Short Period, the b.c. cubic structure is again found in Groups V and VI, but continues into Groups VII (δMn) and VIII (δ, αFe). The f.c. cubic structure is again found in Groups VIIIB and C, but extends backwards into Groups VIIIA (γFe) and VII (γMn). The c.p. hexagonal structure is found in Group VIIIB (Co), but it is probable that this is not strictly comparable with the c.p. hexagonal structures of Ru and Os†.

That Group Numbers VI and VII represent a critical stage in the transition process is suggested by the existence of the numerous allotropes of Mn, and by the formation of characteristic σ structures in alloys of which one metal is of Group Number lower than VII, and the other is of Group Number VII or higher. It is to be noted that the σ structure is also formed by the pure metal βU in Group VI, although here the interpretation is less clear, because of the actinide process.

Recent experimental work has shown that, apart from the σ structures, these alloys may give rise to phases with (*a*) c.p. hexagonal, (*b*) α–Mn and (*c*) β–W types of structure, and that the composition limits of these are often clearly related to the Group Numbers of the constituent elements. The present examination suggests that on proceeding from left to right (i.e. with increasing average Group Number) across the diagram, the different phases occur in the following order:

$$\text{b.c. cube}\rightarrow\beta\text{W}\rightarrow\sigma\rightarrow\alpha\text{Mn}\rightarrow\text{c.p. hex.}\rightarrow\text{f.c. cube.} \quad . \quad . \quad . \quad (1)$$

In all systems one or more of the above phases are absent, but the characteristic order is retained. The general tendencies are seen most clearly by considering the alloys of elements of Groups V and VI in separate diagrams.

1.1. *Alloys of Elements of Group VI*

Figure 1 indicates the approximate composition limits of the different phases in the alloys of Cr, Mo and W with the sequences of elements (Mn, Fe, Co, Ni), (Tc, Ru, Rh, Pd), and (Re, Os, Ir, Pt).

In the alloys of Cr, the σ-phase fields show the characteristic shift from right to left on passing from Mn→Fe→Co, and from Re→Os‡. The c.p. hexagonal or ϵ-phases move in the same direction on passing from Ru→Rh. The β–W phases are at approximately the same composition in the alloys with Ru, Rh, Ir and Pt. In all systems the order of the phases is that of relation (1).

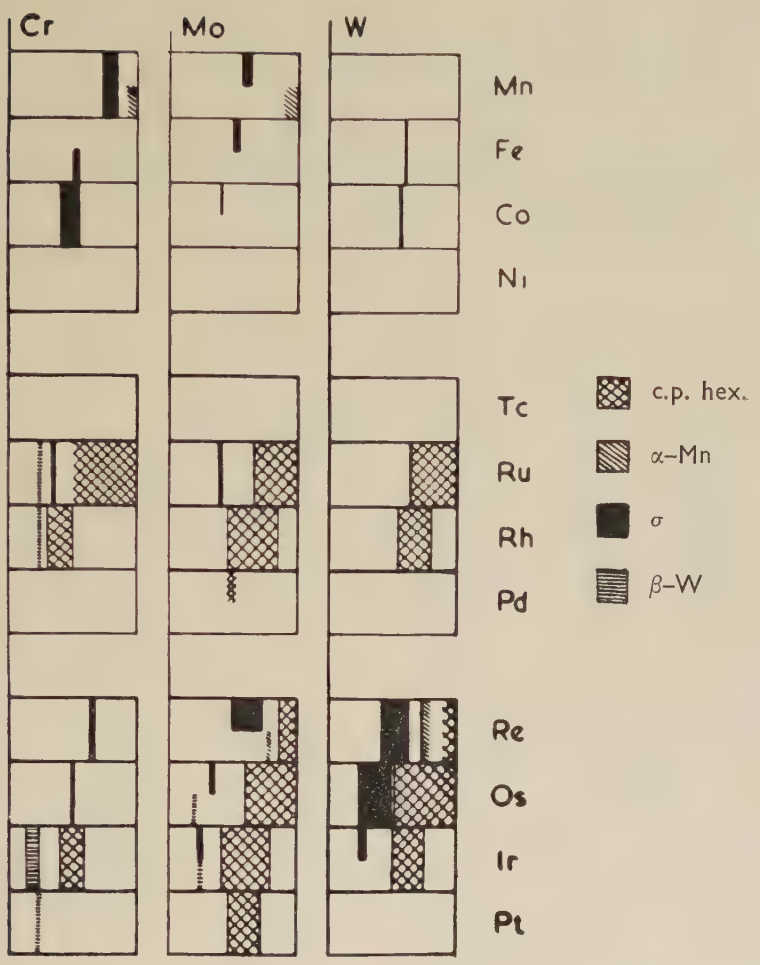
In the alloys of Mo, the σ-phases show the same characteristic shift from right to left in the sequences Mn→Fe→Co, and Re→Os→Ir. The ϵ-phases show the same shift from right to left in the sequences Ru→Rh→Pd, and Re→Os→Ir, although there is little change in the mean composition of the ϵ-phase on passing from Ir→Pt. In all cases the order of the phases is that of relation (1).

---

† The axial ratios for Ru and Os are appreciably smaller than that for c.p spheres in contrast to the value for Co.

‡ This has been pointed out by previous authors.

Fig. 1



In the alloys of W, the σ-phases show the characteristic shift from right to left in the sequences Fe→Co, and Re→Os→Ir. The ϵ-phases show the same shift in the sequences Ru→Rh, and Os→Ir. The order of the phases is always that of relation (1).

1.2. *Alloys of Elements of Group V*

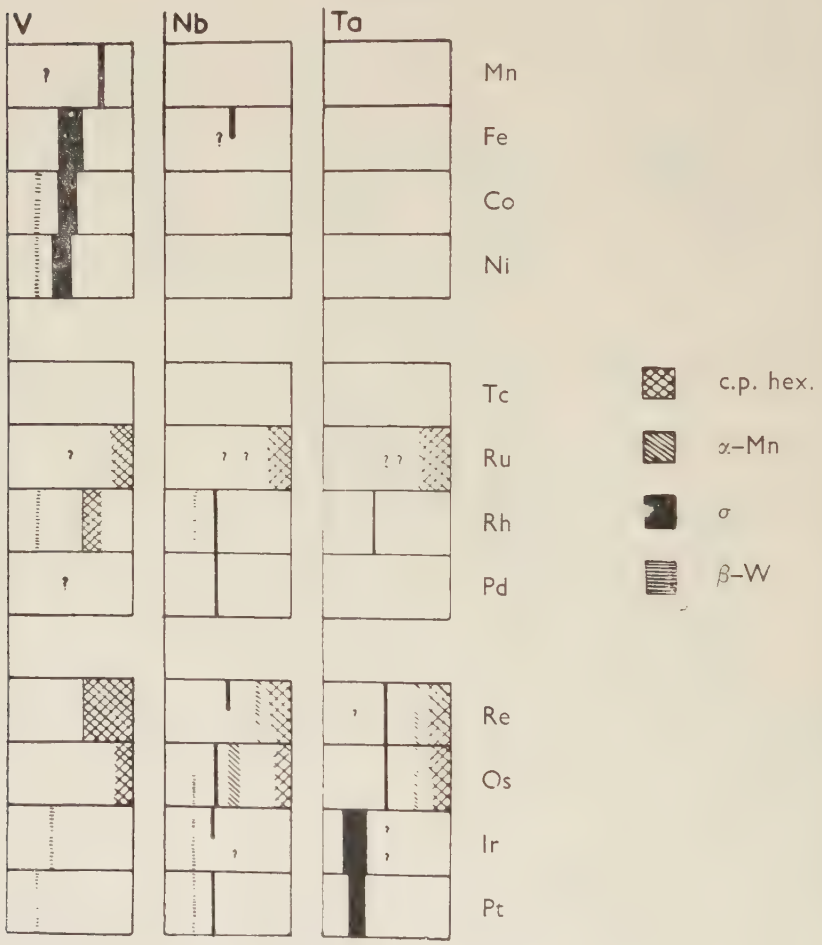
Figure 2 shows the approximate composition limits of the different phases in alloys of V, Nb, and Ta with the same elements as those in fig. 1. In all cases the order of the phases is that of relation (1).

In the alloys of V, the σ-phases show the characteristic shift in the sequence Mn→Fe→Co→Ni, and the ϵ-phases show a shift in the same direction in the sequence Ru→Rh. In the next Period, however, the

solid solution of V in the c.p. hexagonal Re is of wide extent, and on passing from Re→Os the mean composition of the ϵ-phase field may move from left to right.

In the alloys of Nb, the σ-phases show the characteristic shift in the sequence Re→Os→Ir, although there is little change in the compositions of the σ-phases in the sequences Rh→Pd and Ir→Pt. The α–Mn phases show a shift in the same direction in the sequence Re→Os, although the mean composition of the primary ϵ solid solutions may change very slightly in the reverse direction.

Fig. 2



The alloys of Ta are very incompletely known, but the σ-phase fields show the usual shift on passing from (Re, Os)→(Ir, Pt), although there is little change on passing from Re→Os or from Ir→Pt.

From the above survey and from the data of figs. 1 and 2, the following general conclusions may be drawn:

(*a*) The compositions of the βW phases are always very near to the ratio A_3B, where A is the element from Group V or VI, and B is from Group VIII.

(*b*) With very few exceptions, the compositions of the σ, α–Mn, and ϵ-phases, in the alloys of any one metal in Groups V or VI, show a general shift from right to left as the Group Number of the second metal increases from Group VII→VIIIA→VIIIB→VIIIC.

(*c*) The order of the phases given in relation (1) above is always obeyed.

§ 2. Electron Concentration Effects

The changes in composition referred to in (*b*) above are in the direction which would be expected if each of the different structures were favoured by a characteristic electron : atom ratio, where the number of electrons is that outside the inert gas shell. Attempts to correlate the compositions of the σ-phases in this way have been made by several authors, but a detailed examination shows that, although a general correspondence does exist in many cases, it is not possible to account for all the composition ranges in terms of a single electron : atom ratio.

The vacancies in the atomic orbitals postulated by Pauling diminish regularly on passing from Group VI to Group VIIIC, and the correlation in terms of electron : atom ratios is thus accompanied by a correlation in terms of Pauling vacancies, but again the facts cannot all be covered by a simple assumption.

A detailed examination suggests that the composition ranges, not only of the σ-phases, but of all the phases in relation (1) do show a rough correlation with electron atom ratios, and that this is shown most clearly by considering the Group V and Group VI elements separately. The following table refers to alloys of these elements with those of Group VII and VIII. These values must be looked upon as approximate only, because many of the equilibrium diagrams have not been determined accurately.

Table 1

Type of structure	Group V metals	Group VI metals
b.c. cubic	5–5·2	6–6·2
βW	5·8–6·3	6·4–7
σ	6–7·4	6·5–7·6
αMn	6·5–7·2	6·7–7
c.p. hex.	6·2–8	7–8
f.c. cubic	7·5–10	7·5–10

§ 3. The Mutual Solubilities of Transition Metals in the Solid State

Recent work by the authors on the constitution of molybdenum alloys has shown the maximum solubility of molybdenum in palladium to be much greater than that of palladium in molybdenum. Examination of the existing data suggests that this is a general characteristic of the alloys of Group VIIIC with the elements in the middle of the transition series.

Table 2

System	Solubility in Atomic% in element of Group VIIIC	Solubility of Group VIIIC element in second metal
Ni–Cr	46	33
Ni–Ti	13·8	10·3
Ni–Mo	ca 25	small
Ni–Nb	10	<7
Ni–W	17·5	*ca.* 4
Pd–Cr	60	5
Pd–Mo	*ca.* 40	*ca.* 5
Pd–W	*ca.* 20	probably small†
Pt–Cr	75 or 62‡	small

† The lattice spacings of the W-rich phase in 2-phase alloys were indistinguishable from that of pure W.

‡ The value 75 atm% Cr is from the diagram in Vol. 1 of Smithells' Reference Book. The alternative value of 62% is from Raub and Mahler (1955).

Table 2 shows the data for the alloys of elements of Group VIIIC with those of Groups IVA, VA, and VIA, and the general tendency is clear, although the difference in solubilities is less where both elements are in the First Short Period. This tendency would not be expected from the Pauling scheme which allots the same valency of 5·78 to all the metals from Groups VIA to VIIIC. It would, however, agree with the views of Hume-Rothery *et al.* (1951) and of Altmann *et al.* (1957) according to which the valency (in the sense of the number of bonding electrons) diminishes on passing through Group VIII, and the bonding orbitals involve the greatest proportion of *d* function in Group VI. If in Group VIIIC some of the *d* electrons have entered atomic and non-bonding orbitals, it may not be possible for these atoms to give rise to the type of bonding required for the crystal structure of the Group VI element, although there will be no corresponding restriction to prevent the atom of the Group VI element from entering the lattice of an element of Group VIIIC. In this way the marked difference between the extents of the primary solid solutions at the two sides of the equilibrium diagram can be understood, and it should be emphasized that the difference is shown in cases (e.g. Mo–Pd) where the size factor is extremely favourable.

§ 4. Sources of Data

Most of the data on the formation of intermediate phases has been taken from a paper by Greenfield and Beck (1956), and a more recent report from the A.E.I. Laboratories by A. G. Knapton (1957). Reference has also been made to the work of Raub (1954) and Raub and Walter (1951), and to the Metals Reference Book by C. J. Smithells. The data on Mo-Pd is from unpublished work in Oxford by C. W. Haworth and W. Hume-Rothery.

References

Altmann, S. L., Coulson, C. A., and Hume-Rothery, W., 1957, *Proc. roy. Soc.* A, **240,** 145.

Greenfield, P., and Beck, P. A., 1956, *J. Metals*, **8,** 265.

Hume-Rothery, W., Irving, H. M., and Williams, R. J. P., 1951, *Proc. roy. Soc.* A, **208,** 431.

Knapton, A. G., 1957, *A.E.I. Report*, No. A. 682.

Raub, E., 1954, *Z. Metallk.*, **45,** 23.

Raub, E., and Mahler, W., 1955, *Z. Metallk.*, **46,** 210.

Raub, E., and Walter, P., 1951, *W. C. Heraeus Festschrift.*

A New Method for the Evaluation of Electric Conductivity in Metals†

By S. F. Edwards
Department of Mathematical Physics, University of Birmingham

[Received May 23, 1958]

Abstract

A method is developed which allows the evaluation of the closed formal expressions for electrical conductivity which have recently been developed by several authors. The case of a random set of scatterers is treated in detail and the formal solution made to yield directly the solution to the Boltzmann equation. A brief mention of the application of this method to liquids and alloys is made.

§ 1. Introduction

Recently it has been realized by several workers (Nakano 1956, Kubo 1956, Kohn and Luttinger 1957, Greenwood 1958), that the electric conductivity in, say, a metal can be written down in a closed formal expression, without going through the intermediate form of deriving a transport equation, and moreover these closed forms are exact. The usual derivation of a transport equation (cf. Peierls 1955) is rather limited in its applicability and cannot in any simple way be extended to the cases of alloys and liquids etc., and moreover, even where it is usually used, it is not at all clear (see e.g. Peierls 1955, p. 123) that there are not temperature dependent corrections which would entirely invalidate the usual solution. Now the formal exact solutions avoid all this, but carry the difficulty that they are still in a rather abstract form, and it is not clear how they are to be evaluated. This paper is concerned with the evaluation of these formulae, and will show that they can readily give the same result as the usual transport equation where the latter has been assumed to be correct, and thus dispose of the possibility of temperature dependent corrections. The use of the exact formulation in new problems will only be very briefly mentioned in this paper, and since the present object is only to illustrate the method, the simplest problem, that of the conductivity of electrons scattered by a random set of scattering centres, will be discussed.

§ 2. Formulation of the Problem

The starting point will be the formula of Greenwood and Peierls, which states that the conductivity tensor is given by

$$\sigma_{\mu\nu} = -2\pi e^2\hbar \sum_{n,m} v_{nm}{}^{\mu} v_{mn}{}^{\nu} \delta(E_n - E_m) \frac{\partial f}{\partial E_n}, \quad . \quad . \quad . \quad . \quad (1)$$

† Communicated by Professor R. E. Peierls, C.B.E., F.R.S.

where $v_{mn}{}^{\mu}$ is the matrix element of velocity, f the electron distribution function, and the δ function is to be understood in the sense that the limit $E_n \to E_m$ is taken after the system is considered so large that always there are many levels between E_n and E_m. This form as it stands is not suitable for computation, so it is rearranged by first writing it out in full:

$$v_{nm}{}^{\mu} v_{mn}{}^{\nu} = -\frac{\hbar^2}{m^2} \int \psi_n(x) \frac{\partial \psi^*(x)}{\partial x_\mu} d^3x \int \psi_m(y) \frac{\partial \psi_n{}^*(y)}{\partial y^\nu} d^3y. \quad . \quad (2)$$

Introduce units so that $(2m/\hbar^2)^{-1/2} = 1$, then

$$\sum_{m,n} v_{nm}{}^{\mu} v_{mn}{}^{\nu} = -\left(\frac{8\pi e^2}{\hbar}\right) \sum_{n,m} \iiiint \frac{\partial}{\partial x_\mu'} \{\psi_n(x)\psi_n{}^*(x')\} \frac{\partial}{\partial y_\nu'} \{\psi_m(y)\psi_m(y')\}$$
$$\times \delta(x'-y)\delta(y'-x)\, d^3x\, d^3y\, d^3x'\, d^3y'. \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Now $\sum_n \psi_n(x)\psi_n(x')\delta(E-E_n)$ is a Green function, the solution of the homogeneous Schrödinger equation. If the Schrödinger equation for ψ is

$$(H-E)\psi = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and the Green functions G_+, G_- are given by

$$\left.\begin{aligned} (H-E+i\epsilon)G_+(x,x') &= \delta(x-x') \\ (H-E-i\epsilon)G_-(x,x') &= \delta(x-x') \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where ϵ is an infinitesimal quantity used to define the contour defining the G_+, G_-, then

$$G_+(x,x') = \sum_n \psi_n(x)\psi_n{}^*(x')(E-E_n+i\epsilon)^{-1} \quad . \quad . \quad . \quad . \quad (6)$$

$$G_-(x,x') = \sum_n \psi_n(x)\psi_n{}^*(x')(E-E_n-i\epsilon)^{-1}. \quad . \quad . \quad . \quad . \quad (7)$$

From these the sum and the difference can be made

$$G_+ - G_- = 2\pi i \sum \psi_n(x)\psi_n{}^*(x')\delta(E-E_n)$$
$$= 2\pi i G \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$G_+ + G_- = 2\pi i P \sum \psi_n(x)\psi_n{}^*(x')(E-E_n)^{-1}$$
$$= 2\pi i G_P \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

P standing for principal part.

In the absence of potentials these functions are just

$$(4\pi r)^{-1} e^{i\sqrt{(Er)}}, \qquad (4\pi r)^{-1} e^{-i\sqrt{(Er)}}, \qquad i(4\pi^2 r)^{-1} \sin\sqrt{(Er)},$$
$$i(4\pi^2 r)^{-1} \cos\sqrt{(Er)}, \qquad (r = |\mathbf{X}-\mathbf{X}'|), \quad . \quad . \quad . \quad . \quad (10)$$

assuming that one already is dealing with an infinite system, i.e., a continuum of energy so that the sums over n become integrals. So if G is used for the difference of G_+ and G_-, (8) the sine like function, and also

we now specialize to the case of a diagonal σ, we have

$$\sigma = \frac{1}{3}\left(\frac{8\pi e^2}{\hbar}\right)\sum_{n,m}\iiiint (\partial G_n(x,x')/\partial x_\mu)(\partial G_m(y,y')/\partial y_\mu) \times \delta(E_n - E_m)(\partial f/\partial E_n)\delta(x'-y)\partial(y'-x)\, d^3x\, d^3y\, d^3x'\, d^3y' \quad . \quad (11)$$

where $G_n(x,x')$ is $G(E_n; x, x')$. The problem now is to find G and f in the presence of the scattering potential, and finally to average σ over all configurations of this potential†. It is convenient to express this in the following way: the Schrödinger equation is now

$$(\nabla^2 - E \pm i\epsilon + V(\mathbf{x}))G_\pm(x,x') = \delta(x-x'), \quad . \quad . \quad . \quad (12)$$

where

$$V(\mathbf{x}) = \sum_\alpha u(\mathbf{x} - \mathbf{X}_\alpha) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

X_α being the positions of the scattering centres, and u the potential they exert on the electron. It is convenient to use the Fourier transform of this potential, defining

$$\rho_{\mathbf{k}} = \sum_\alpha e^{i\mathbf{k}\mathbf{X}_\alpha}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

$$V(x) = \int u(k)\, e^{i\mathbf{k}\mathbf{x}}\rho_{\mathbf{k}}^*\, d^3k. \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

Now if the $\mathbf{X}_\alpha$ are random, it can be shown by standard probability theory that the distribution function for the ρ's is

$$\left.\begin{aligned} P(\rho_{\mathbf{k}}) &= \xi \exp\left[- \iint R(\mathbf{k},\mathbf{j})\rho_{\mathbf{k}}\rho_{\mathbf{j}}\, d^3k\, d^3j \right. \\ &\qquad \left. - \iiiint Q(\mathbf{k},\mathbf{j},\mathbf{l},\mathbf{m})\rho_{\mathbf{k}}\rho_{\mathbf{j}}\rho_{\mathbf{l}}\rho_{\mathbf{m}}\, d^3k\, d^3j\, d^3l\, d^3m - \ldots \right] \\ R(\mathbf{k},\mathbf{j}) &= N^{-1}\delta(\mathbf{k}+\mathbf{j}) + O(N^{-3}) \\ Q(\mathbf{k},\mathbf{j},\mathbf{l},\mathbf{m}) &= N^{-3}\left(\frac{1}{24}\right)\left[\delta(\mathbf{k}+\mathbf{j}+\mathbf{l}+\mathbf{m}) - 3\sum_{\text{perm}}\delta(\mathbf{k}+\mathbf{j})\delta(\mathbf{l}+\mathbf{m})\right] + O(N^{-5}) \end{aligned}\right\}, \quad . \quad . \quad . \quad (16)$$

where N is the total number of scatterers, and ξ the normalization to give total probability unity. When N is large, this can be used with k running over the whole continuum of k space. So we reach the final formula

$$\sigma = \frac{1}{3}\left(\frac{8\pi e^2}{\hbar}\right)\int \ldots \int \xi P(\rho_{\mathbf{k}})(\partial G_n(x,x')/\partial x_\mu)(\partial G_m(y,y')/\partial y_\mu) \times \delta(E_n - E_m)(\partial f/\partial E_n)\delta(x-y')\delta(y-x')\, d^3x\, d^3y\, d^3x'\, d^3y' \prod d\rho_{\mathbf{k}}\, d\rho_{\mathbf{k}}^*. \quad (17)$$

This form has the great advantage that it is essentially the same form as that of electrons interacting with the quantized electromagnetic field, and so techniques for evaluating it are already in existence. Moreover, there are none of the divergence problems of electrodynamics here and the various approximate techniques of electrodynamics can be applied with confidence.

† The meaning of the averaging is discussed in detail by Kohn and Luttinger (1957).

§ 3. Evaluation

The essential difference between (1) and (17) is that the averaging over the scatterers can be carried out before the integrations over coordinates and allows manipulations which are meaningless when applied to (1). Although one can, on the basis of (17), derive integral equations for the average of $G(x,x')$, $G(y,y')$, it is simplest to consider the perturbation expansion of the G's, from which the structure of the integral will become clear. However, one should emphasize that there is no need to approach the evaluation by a perturbation approach and the results to be obtained below can be got directly.

Consider firstly a simpler problem, that of obtaining the average of just one G alone. This is the difference of the averages of G_+ and G_-, which are more convenient to consider. Now in perturbation theory one can write, using $G^{(0)}$ for the ρ-independent functions (10),

$$G_+(x,x') = G_+^{(0)}(x,x') - \iint G_+^{(0)}(x,y)u(\mathbf{k})\,e^{i\mathbf{k}\mathbf{y}}\rho_{\mathbf{k}}{}^*G_+^{(0)}(y,x')\,d^3y\,d^3k$$

$$+ \iiiint G_+^{(0)}(x,y)u(\mathbf{k})\,e^{i\mathbf{k}\mathbf{y}}\rho_{\mathbf{k}}{}^*G_+^{(0)}(y,z)u(\mathbf{j})\,e^{i\mathbf{j}\mathbf{z}}\rho_{\mathbf{j}}{}^*G_+^{(0)}(z,x')$$

$$\times d^3y\,d^3z\,d^3k\,d^3j + \ldots . \quad . \quad . \quad (18)$$

Upon averaging, using brackets for average value

$$\langle \rho_{\mathbf{k}} \rangle = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

$$\langle \rho_{\mathbf{k}}\rho_{\mathbf{j}} \rangle = N\delta(\mathbf{k}+\mathbf{j}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

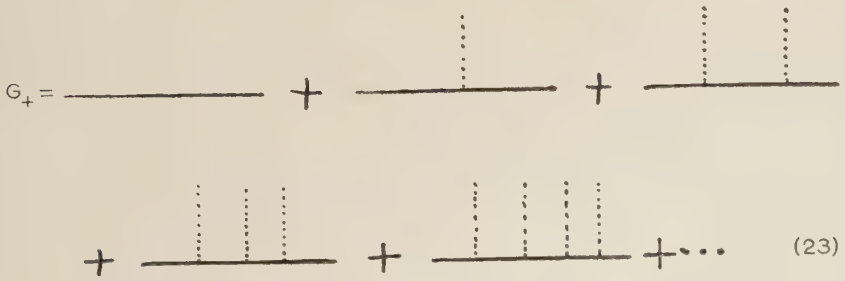
$$\langle \rho_{\mathbf{k}}\rho_{\mathbf{j}}\rho_{\mathbf{l}}\rho_{\mathbf{m}} \rangle = N^2 \sum_{\text{perm}} \delta(\mathbf{k}+\mathbf{j})\delta(\mathbf{l}+\mathbf{m}) + N\delta(\mathbf{k}+\mathbf{j}+\mathbf{l}+\mathbf{m}) \quad . \quad . \quad (21)$$

and so on, neglecting terms relatively of order N^{-1}. This can be obtained directly of course, without using the expansion (16). This gives

$$\langle G_+(x,x') \rangle = G_+^{(0)}(x,x') + \iiint NG_+^{(0)}(x,y)\,e^{i\mathbf{k}(\mathbf{y}-\mathbf{z})}u^2(\mathbf{k})$$

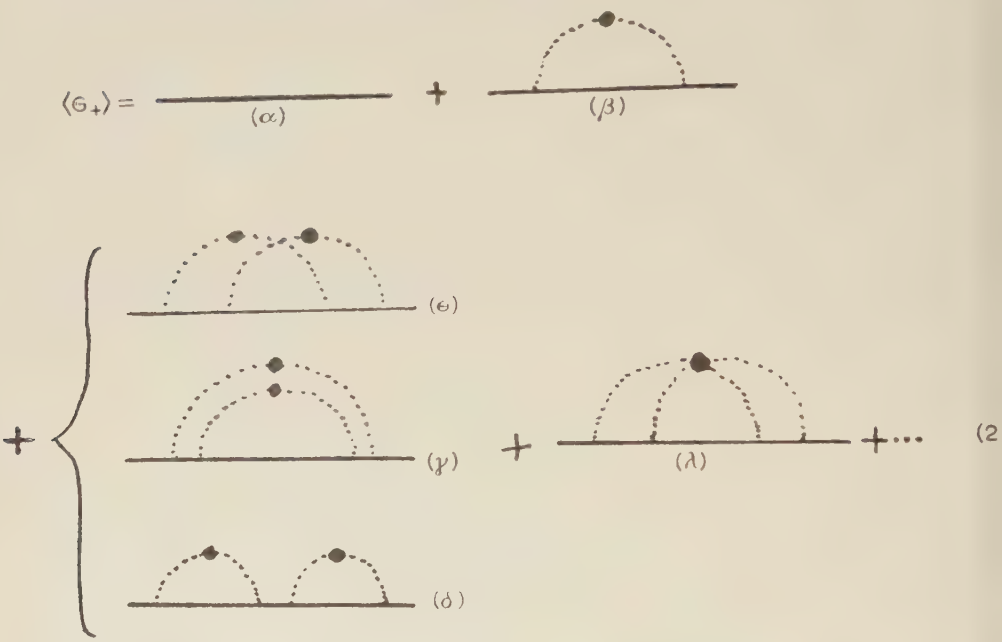
$$\times G_+^{(0)}(y,z)G_+^{(0)}(z,x')\,d^3y\,d^3z\,d^3k + \ldots . \quad . \quad . \quad (22)$$

This is conveniently expressed in diagrams, which are slightly different from those of electrodynamics. Consider $G(x,x')$ before averaging, draw a full line for every $G_+^{(0)}$ and a dotted line for every $u\rho$. Then the expansion of $G(x,x')$ is written



(23)

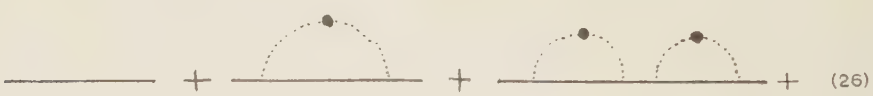
The averaging process joins the loose dotted line ends up, and places the value N at a join of two, N at a join of four, and so on. Mark these joins by a large dot. Then the average appears as



These diagrams can readily be labelled in configuration or momentum space

(25)

the dot having significance only when diagrams like the last in eqn. (24) appear which incidentally has no analogue in electrodynamics†, which is effectively the case of N infinite. Now concentrate attention upon the series



(26)

This is a simple expansion of the series

$$\left(G_+^{(0)-1}(p) - N\int G_+^{(0)}(q)u^2(\mathbf{p}-\mathbf{q})\,d^3q\right)^{-1}. \quad . \quad . \quad . \quad (27)$$

† In electrodynamic language the G_+ is S_F, the dotted line the vertex Γ, the dot the photon Green function D_F.

Since $G_+^{(0)}(q) = (2\pi)^{-3}(q^2 - E + i\epsilon)^{-1}$ the integral can be written

$$\tfrac{1}{2}P\int G_+(q)u^2(p-q)\,d^3q + \frac{i\pi}{(2\pi)^3}\int \delta(q^2-E)u^2(p-q)\,d^3q. \qquad (28)$$

If now u is taken as isotropic so that u^2 is $u^2(p^2+q^2-2pq\cos\theta)$ this can be written as $A+iB$, where A, B are real

$$A = \tfrac{1}{2}NP\int (q^2-E)^{-1}u^2(p^2+q^2-2pq\cos\theta)\,d^3q \quad . \quad . \quad . \quad (29)$$

$$B = (2\pi)^{-3}\pi N\int \delta(q^2-E)u^2(p^2+q^2-2pq\cos\theta)\,d^3q. \quad . \quad . \quad (30)$$

These can be expanded as series in p^2-E so that this approximation to $\langle G\rangle^{-1}$ is in the form

$$(p^2-E+i\epsilon+a_1+ib_1+(p^2-E)(a_2+ib_2)+\ldots) \quad . \quad . \quad . \quad (31)$$

where

$$a_1 = \tfrac{1}{2}(2\pi)^{-3}NP\int (q^2-E)^{-1}u^2(E+q^2-2q\sqrt{(E)}\cos\theta)\delta(p^2-E)\,d^3q\,d^3p \quad (32)$$

$$b_1 = N\sqrt{E}(8\pi)^{-1}\int \sin\theta\,d\theta\,u^2(2E(1-\cos\theta)). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

u^2 is in fact the differential scattering probability in Born approximation, $w_0(\theta) = (4\pi^2)^2 2\pi\hbar^{-1}u^2(2E(1-\cos\theta))$

$$b_1 = (2\pi)^{-6}\sqrt{(E)}\hbar N\int \sin\theta\,w_0(\theta)\,d\theta. \quad . \quad . \quad . \quad . \quad (34)$$

A rather more refined treatment is to expand not in terms of p^2-E but $p^2-E+a_1+ib_1$ an important step in electrodynamics where only a_1-E is defined, but here it makes little odds as we are anyway taking all the a's to be small, and the effect of the terms a_2, b_2 etc. will come out very small. So it has been found that this series summed gives effectively a complex displacement of the energy E

$$\langle G_+(E)\rangle \cong G_+(E+\delta E) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

where

$$\begin{aligned}(E+\delta E)^{1/2} &= (E+a_1+ib_1)^{1/2}\\ &\cong \sqrt{E}+(a_1+ib_1)/2\sqrt{E}\\ &= \sqrt{E'}+i\Gamma. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)\end{aligned}$$

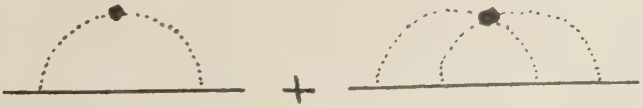
So in configuration space

$$\langle G_+\rangle = (4\pi r)^{-1}\,\mathrm{e}^{-i\sqrt{(E')}r-\Gamma r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

(cf. Bardeen 1956). If this calculation had been performed for G_- the result would have been

$$\langle G_-\rangle = (4\pi r)^{-1}\,\mathrm{e}^{i\sqrt{(E')}r-\Gamma r}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

Now consider to what extent one can take the forms (26), (37) as adequate approximations to the whole series. Consider first those terms containing one dot only, in particular the series



(39)

This series is in fact building up the exact scattering of one electron by one scattering centre, instead of its first Born approximation, and if instead of taking the unit

in the series (26) one took all one dot diagrams one would just replace the Born approximation scattering by the true differential scattering cross section in (34), $w(\theta)$. This can be important in strong interaction, and this way of looking at it will be valuable if one can think of the electron in strong interaction with the scattering centres one at a time, as is usually considered to be the case. In a dense system, however, the electron interacts with many centres at once, and one cannot disentangle the scattering with one centre from that of all the others, i.e. the other terms to be discussed below. Henceforward these diagrams will be ignored except inasmuch as the differential scattering cross section can be understood as the true one rather than Born approximation. Now turn to the other terms in (24), in particular, say, (δ) and (ϵ). The electrical conductivity based on the approximate sum (31) and further calculations to be given below comes out to be of order Γ^{-1}, i.e. inversely with the square of the interaction. Terms like (δ), (ϵ) and the higher terms, if included, give a series in the interaction, but do not alter the first term which will still dominate the calculation. To see this it is perhaps simplest to look at G_+ in configuration space. The series which has been considered so far (26) amounts to

$$(4\pi r)^{-1}\,\mathrm{e}^{i\sqrt{(E')}r-\Gamma r}=(4\pi r)^{-1}\,\mathrm{e}^{i\sqrt{(E)}r}(1+Lr+\tfrac{1}{2}L^2r^2+\ldots) \tag{40}$$

where

$$L=i(\sqrt{E'}-\sqrt{E})-\Gamma. \tag{41}$$

The inclusion of terms like (δ), (ϵ) etc. adds in terms so that in first order Lr is corrected by a constant, in second order $\frac{1}{2}L^2r^2$ is corrected by a term in r, and so on, always a power of r less in any order, so that summing on the basis of (31) one has

$$\langle G_+\rangle=(4\pi r)^{-1}\,\mathrm{e}^{(i\sqrt{E'}-\Gamma)r}(1+k_0r+k_1r^2+\ldots) \tag{42}$$

and, as will appear below, this affects the conductivity by second and higher terms in u^2

$$\sigma\sim O(\Gamma^{-1})+O(1)+O(\Gamma)+O(\Gamma^2)+\ldots. \tag{43}$$

Thus (37) is a good basis for the evaluation of $\langle G_+\rangle$ and hence, of course, $\langle G\rangle$.

To summarize in a rather formal way the above discussion, consider the identity

$$\begin{aligned}G_+(x,x')=G_+^{(0)}(x,x')-\iint G_+^{(0)}(x,y)\,\mathrm{e}^{i\mathbf{k}\mathbf{y}}u(\mathbf{k})\rho_{\mathbf{k}}G_+^{(0)}(y,x')\,d^3y\,d^3k\\+\iiiint G_+^{(0)}(x,y)\,\mathrm{e}^{ik\mathbf{y}}u(\mathbf{k})\rho_{\mathbf{k}}G_+^{(0)}(y,z)u(\mathbf{j})\,\mathrm{e}^{i\mathbf{j}\mathbf{z}}G_+(z,x')\\\times d^3y\,d^3z\,d^3k\,d^3j.\end{aligned} \tag{44}$$

Then

$$\langle G_+(x,x')\rangle = G_+^{(0)}(x,x') - \iint G_+^{(0)}(x,y)\, e^{iky} u(\mathbf{k}) \langle \rho_{\mathbf{k}} \rangle G_+^{(0)}(y,x')\, d^3y\, d^3k$$
$$+ \iiiint G_+^{(0)}(x,y)\, e^{i\mathbf{k}\mathbf{y}} u(\mathbf{k}) \langle \rho_{\mathbf{k}} G_+^{(0)}(y,z) u(\mathbf{j})\, e^{i\mathbf{j}\mathbf{z}} \rho_{\mathbf{j}} G_+(z,x')\rangle$$
$$\times d^3y\, d^3z\, d^3k\, d^3j \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

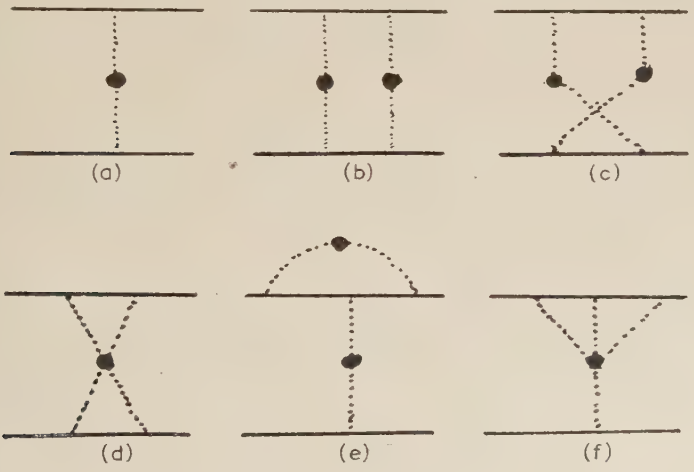
$$= G_+^{(0)}(x,x') + \iiiint G_+^{(0)}(x,y) v(\mathbf{k}) v(\mathbf{j})\, e^{i\mathbf{k}\mathbf{y} + i\mathbf{j}\mathbf{z}} G_+^{(0)}(y,z)$$
$$\times \langle \rho_{\mathbf{k}} \rho_{\mathbf{j}} G_+(z,x')\rangle\, d^3y\, d^3z\, d^3k\, d^3j \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (46)$$

$$= G_+^{(0)}(x,x') - \iint G_+^{(0)}(x,y) \Sigma(y,z) \langle G_+(z,x')\rangle\, d^3y\, d^3z, \quad . \quad . \quad (47)$$

which defines $\Sigma(y,z)$. Then symbolically

$$[(G_+^{(0)})^{-1} + \Sigma]\langle G_+\rangle = \delta. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (48)$$

The discussion above has shown that when the interaction is weak, an evaluation of Σ by perturbation theory is adequate. So far all the discussion has concerned $\langle G_+\rangle$, but (17) involves the evaluation of $\langle G_n' G_m' \partial f/\partial E\rangle$, and the two G's and the f will interfere. The dependence of f upon the p's in the averaging does not affect the answer in its leading term, so it is ignored for the present. Since the problem is already being considered for an infinitely large conductor the distinction between n and m can be dropped at this stage. Consider at first the quantity $\langle G(x,x')G(y,y')\rangle$. In terms of the diagrams, this consists of two full lines, with dotted lines leaving and entering both, in particular, in addition to the types of series (24) one also has types



(49)

Types (*d*), (*f*) clearly belong to the same category as those of (24, λ) and will be considered no further. Also disregard type (*e*) which is an interference between type (α) and (β), and corresponds to the Lamb shift in electrodynamics: it can readily be found to be small when the interaction is weak. Of the remainder, types (*a*), (*b*) are again the first two of a simple

series analogous to (β), (γ) of (24). Assuming then that terms like (a), (b) do not interfere with terms like (β), (γ) one has the equation for $\langle GG\rangle$:

$$\langle G(x,x')G(y,y')\rangle = \langle G(x,x')\rangle\langle G(y,y')\rangle + N\iiint \langle G(x,z)\rangle\langle G(y,w)\rangle u^2(\mathbf{k}) \times e^{i\mathbf{k}(\mathbf{z}-\mathbf{w})}\langle G(z,x')G(w,y')\rangle\, d^3k\, d^3z\, d^3w, \quad . \quad (50)$$

or in the spirit of (48) the exact equation can be written symbolically in terms of a generalized 'interaction' I

$$\langle GG\rangle = \langle G\rangle\langle G\rangle + \langle G\rangle\langle G\rangle I\langle GG\rangle \quad . \quad . \quad . \quad . \quad (51)$$

and as in (48), by explicit calculation of the errors, one finds that I can be evaluated by perturbation theory if the interaction is weak. The first approximation, the sum of (a), (b) . . ., is then (51) with

$$I(\alpha,\beta) = N\int d^3k u^2(\mathbf{k})\, e^{i\mathbf{k}(\alpha-\beta)}. \quad . \quad . \quad . \quad . \quad . \quad (52)$$

The quantity required has $x'=y$, $y'=x$, so let Δ be $\langle GG\rangle$ in this case. Then if Δ is written in momentum space, the quantity required in (17) is, where Ω is the total volume,

$$\Omega\int \mathbf{p}\cdot\mathbf{q}\,\Delta(\mathbf{p},\mathbf{q})\, d^3p\, d^3q \quad . \quad . \quad . \quad . \quad . \quad . \quad (53)$$

where (52) becomes, putting

$$\langle G(p)\rangle\langle G(-p)\rangle = g(p) \quad . \quad . \quad . \quad . \quad . \quad . \quad (54)$$

$$\Delta(\mathbf{p},\mathbf{q}) = g(p)\delta(\mathbf{p}-\mathbf{q}) + N\int g(p)u^2(\mathbf{p}-\mathbf{s})\Delta(\mathbf{s},\mathbf{q})\, d^3s. \quad . \quad . \quad (55)$$

Let

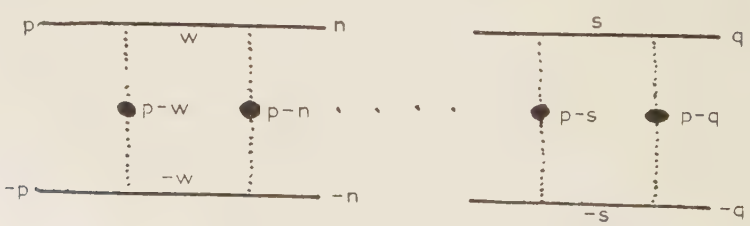
$$\int \mathbf{p}\cdot\mathbf{q}\,\Delta(\mathbf{p},\mathbf{q})\, d^3q = K(p) \quad . \quad . \quad . \quad . \quad . \quad . \quad (56)$$

then

$$K(p) = p^2g(p) + N\int g(p)u^2(\mathbf{p}-\mathbf{s})\mathbf{p}\cdot\mathbf{q}\,\Delta(\mathbf{s},\mathbf{q})\, d^3s\, d^3q \quad . \quad . \quad . \quad (57)$$

$$= p^2g(p) + 2\pi N\int g(p)u^2(p^2+\delta^2-2ps\cos\theta)\cos\theta\sin\theta \times K(s)s^2\, ds\, d\theta. \quad . \quad . \quad . \quad . \quad . \quad (58)$$

The fact that $\mathbf{p}\cdot\mathbf{q} = pq\cos\theta_{pq}$ effectively contributes $pq\cos\theta_{ps}$ is seen by writing all the quantities concerned in spherical harmonics, or perhaps simpler by considering the series of which (52) is the sum, consisting of terms like



The integration over $\mathbf{q}$ with $\cos\theta_{pq}$ leaves a $\cos\theta_{ps}$ for the s integration, and so right through the diagram, which is in effect (58).

Now $g(p)$, apart from the displacement Γ is just a function with $p=\sqrt{E}$ or more accurately $p=\sqrt{E'}$. $K(p)$ has essentially this same feature, so to the accuracy that has been used so far, the solution of (58) can be written down at once

$$K(p)=p^2g(p)\left[1-2\pi N\int u^2(2E(1-\cos\theta))\cos\theta\sin\theta\int q(s)\delta^2\,ds\right]^{-1}. \qquad (59)$$

The integral of K can be obtained by contour integration or by the discussion below in equations (64)–(66).

$$\int K(p)\,d^3p=\frac{E'/2\Gamma'}{1-\Gamma'/\Gamma} \qquad (60)$$

$$=\frac{E'}{2}(\Gamma-\Gamma')^{-1} \qquad (61)$$

where

$$\Gamma'=(8\pi)^{-1}N\hbar\int u^2(2E'(1-\cos\theta))\sin\theta\cos\theta\,d\theta \qquad (62)$$

or in the notation of (34)

$$\Gamma'=\tfrac{1}{2}(2\pi)^{-6}N\hbar\int\cos\theta\sin\theta\,w_0(\theta)\,d\theta. \qquad (63)$$

It is perhaps useful to look at (60) in configuration space. Without scattering the conductivity is

$$\sigma=\frac{1}{3}\left(\frac{8\pi e^2}{\hbar}\right)\frac{\Omega}{(2\pi)^4}\sum_E\int\nabla\frac{\sin\sqrt{E}r}{r}\nabla\frac{\sin\sqrt{E}2}{r}(\partial f/\partial E)\,d^3r. \qquad (64)$$

This diverges at large distances where the integral over r looks like

$$\sim\int^{\infty}dr.$$

When the scattering is introduced, the term $\sin\sqrt{(E)}r$ is replaced by $\sin\sqrt{(E')}re^{-\Gamma r}$ and also the averaging process introduces the cross term $e^{\Gamma' r}$ giving altogether, at large distances,

$$\int e^{-2(\Gamma-\Gamma')r}\,dr \qquad (65)$$

$$\sigma=\Omega e^2/(12\pi^2\hbar(\Gamma-\Gamma')). \qquad (66)$$

This, of course, is the usual answer (cf. Peierls 1955, eqns. (6.16), (6.20)); Peierls w is our Nw/Ω, in particular for free electrons where $\partial f/\partial E$ is approximately a δ function at the Fermi surface, so integration over the Fermi surface gives the final result

$$\sigma=e^2n_e\left(4\pi mn_s\int w(\theta)(1-\cos\theta)\sin\theta\,d\theta\right)^{-1} \qquad (67)$$

where n_e is the density of electrons, and n_s the density of the scatterers.

The dependence of f upon the configuration of scatterers can be expressed by expanding it in a power series in the density of scatterers and adding in the terms so produced. These corrections are of the same order as those in (43), and since they are well defined at $T=0$, do not involve the temperature in any critical form.

It is worth remarking that in averaging over all configurations, one includes those configurations for which the conductivity is infinite, for example an ordered lattice system. The method of calculating presented here automatically gives these configurations negligible weight, but a rigorous treatment would require a more careful treatment.

§ 4. Discussion

It has been shown that the exact formal solution of the equation of motion can be evaluated to give the usual solution of the Boltzmann transport equation, and within the framework of weak interaction it is quite easy to write down higher order corrections, though of course these rapidly become very numerous. Of more interest is the possibility of evaluating formula (11) in cases where perturbation theory is not applicable. An example, which is still far from being the most general state of affairs but is of physical interest, is the case when the distribution of scatterers (which may be lattice vibrations, etc.) is known through a partition function, and the electrons still interact weakly with the scattering centres. This is a model of a liquid or an alloy. For a liquid, eqns. (20), (21) and so on, are not satisfied, and the averages can only be found from the partition function. If it could be assumed to a reasonable degree of approximation that

$$\langle \rho_{\mathbf{k}} \rho_{\mathbf{j}}^* \rho_{\mathbf{l}} \rho_{\mathbf{m}}^* \rangle = \sum_{\text{perm}} F(\mathbf{k}-\mathbf{j}) F(\mathbf{l}-\mathbf{m}) \quad . \quad . \quad . \quad . \quad (68)$$

where

$$F(\mathbf{k}-\mathbf{j}) = \langle \rho_{\mathbf{k}} \rho_{\mathbf{j}}^* \rangle \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (69)$$

then one could immediately write down the conductivity by replacing

$$\int w(\theta)(1-\cos\theta)\sin\theta \, d\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (70)$$

by

$$\int F(2E(1-\cos\theta)) w(\theta)(1-\cos\theta)\sin\theta \, d\theta \quad . \quad . \quad . \quad (71)$$

in (67). Since one is dealing with smooth averages this approximation may be adequate for this problem, even if it is not so for the theory of liquids as a whole. In general, however, the conductivity will involve not only the two body correlation function F, which is available experimentally, but the whole partition function, which at present is not available. There are models available for alloys, however, in particular for super lattice forming alloys and this problem is being considered further. Methods are in existence for evaluation of formulae like (17) in cases where perturbation methods are inapplicable, but a discussion of these will be left until they have been successfully applied.

Acknowledgments

The author would like to thank Professor Peierls for suggesting this problem and for helpful discussions during its solution. He would also like to thank Drs. G. V. Chester and A. Thellung, who have also considered the evaluation of the exact formulation from a different viewpoint, for helpful discussions, and Professor J. M. Luttinger for pointing out some inadequacies of the first version of the work and for drawing the author's attention to the work of Bardeen.

References

Bardeen, J., 1956, *Handbuch der Physik, Berlin*, **15,** 274.
Greenwood, D. A., 1958, *Proc. phys. Soc. Lond.*, **71,** 585.
Kohn, W., and Luttinger, J. M., 1957, *Phys. Rev.*, **108,** 590.
Kubo, R., 1956, *Canad. J. Phys.*, **34,** 1274.
Nakano, H., 1956, *Progr. theor. Phys.*, **15,** 77.
Peierls, R. E., 1955, *Quantum Theory of Solids* (Oxford).

The Deformation of Metals by Self-Diffusion†

By A. P. Greenough
University College, Swansea, Glam.

[Received June 2, 1958]

Abstract

The experimental data on the deformation of metals at temperatures near their melting points is reassessed. It seems that at any given stress, an initial period in which the strain rate decreases is followed by a second period in which the strain rate is constant. The Nabarro–Herring theory satisfactorily accounts for the constant strain rate. In the initial period, the higher strain rate cannot entirely be attributed to offsetting and kinking at grain boundaries. It is suggested that edge dislocations may rotate about suitable nodes, and in planes normal to their slip planes, thus acting as sources or sinks for vacancies during the initial period.

§ 1. Introduction

It is about ten years since Nabarro (1948) and Herring (1950) independently suggested that metals might be able to deform under very low stresses by a self-diffusion mechanism. Essentially, it is postulated that vacant lattice sites diffuse through the crystals, thus effectively producing mass transfer in the opposite direction to the vacancy flow. Both authors considered that the grain boundaries and the free surfaces of the metal specimens could act as sources or sinks for vacancies.

Since these papers were published, the results of a considerable amount of experimental work have appeared. Unfortunately, in many cases, the analysis of the results of this work is incomplete or inaccurate, and therefore it seems appropriate that the work in this field should be critically reviewed.

Herring derives expressions for the relationship between the applied stress (which must be corrected for surface energy effects) and the strain rate, for both wire and foil specimens. In the case of a wire which has a 'bamboo' grain structure, i.e. the grains occupy the whole cross section of the wire, and the grain boundaries are perpendicular to the wire axis, we obtain from his eqn. (16) the expression:

$$\text{Strain rate,} \qquad \frac{1}{L}\frac{dL}{dt} = \frac{1}{L}\frac{D\Omega}{2rRT}\,BS \qquad (1)$$

where L = average length of crystal grains in wire, t = time, D = volume self-diffusion coefficient, Ω = gm atomic volume, r = radius of wire, R = gas constant, T = temperature, B = numerical constant = 12 (approx.) when $L = 2r$, and S = applied stress, corrected for surface energy effects.

† Communicated by the Author.

Both Herring and Nabarro give good reasons why grain boundaries separating grains with less than a minimum orientation difference should not act as sources or sinks for vacancies under these conditions. The strain rate : stress graph for wires of given structure should thus be linear, and the slope of the graph at any given temperature (ϵ_T) may be calculated if D is known.

Since the D may be expressed as follows:

$$D = A \exp(-E/RT) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where A = constant and E = activation energy for self diffusion, E may be found from the gradient of the $\log \epsilon_T/T : 1/T$ graph. For typical metals this is 5–10% higher than the activation energy for creep calculated from the slope of the $\log \epsilon_T : 1/T$ graph.

For a given grain structure, the strain rate at any temperature is directly proportional to the applied stress, and the deformation resembles viscous flow in this respect. For this reason, some authors prefer to compare experiment with theory in terms of apparent viscosity.

§ 2. Review of Experimental Work

2.1. *Copper*

Herring (1950) first analysed the results for copper wires obtained by Udin *et al.* (1949) and Udin (1951). Table 1 compares the gradients of the strain rate : stress graphs obtained experimentally with the gradients calculated from the most recent values of the self-diffusion coefficients of copper obtained by the radioactive tracer technique, using single crystal specimens (Kuper *et al.* 1954). The slopes calculated from the self-diffusion data used by Herring are also shown.

The activation energy for self-diffusion derived from the deformation of the wires is 61 000 cal/mole, compared with 47 140 cal/mole obtained by

Table 1

Temperature °C	Wire diam. cm	Slope of strain rate : stress graph cm^2/dyne/sec		
		Observed	Calculated Kuper *et al.*	Calculated Herring
1050	0·0064	$2{\cdot}70 \times 10^{-13}$	$3{\cdot}60 \times 10^{-14}$	—
1025	0·0064	$2{\cdot}43 \times 10^{-13}$	$2{\cdot}60 \times 10^{-14}$	—
1000	0·0064	$1{\cdot}26 \times 10^{-13}$	$1{\cdot}85 \times 10^{-14}$	$1{\cdot}23 \times 10^{-14}$
950	0·0064	$3{\cdot}38 \times 10^{-14}$	$8{\cdot}98 \times 10^{-15}$	—
1050	0·0036	$1{\cdot}39 \times 10^{-13}$	$5{\cdot}68 \times 10^{-14}$	—
1000	0·0036	$8{\cdot}99 \times 10^{-14}$	$2{\cdot}92 \times 10^{-14}$	$1{\cdot}92 \times 10^{-14}$
950	0·0036	$2{\cdot}75 \times 10^{-14}$	$1{\cdot}42 \times 10^{-14}$	—

the radioactive tracer technique. Contrary to theoretical predictions, the slopes of the strain rate : stress graphs at a given temperature are less for the smaller diameter wires than for the larger diameter wires.

Pranatis and Pound (1955) studied the deformation of copper foil. Unfortunately, they consider that "by definition, the reciprocals of the slopes of the strain rate : stress curves are the viscosities, that is $\sigma = \eta\dot{\epsilon}$", and thus they omit a numerical constant in deriving the apparent viscosity from their experimental measurements.

Frenkel (1945) showed that when a longitudinal force F acts on a viscous rod length L, volume V, the extension of the rod is given by

$$\frac{1}{L^2}\frac{dL}{dt} = \frac{F}{3\eta V} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where η is the viscosity, i.e.

$$\sigma = 3\eta\dot{\epsilon} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where σ = applied stress and $\dot{\epsilon}$ = strain rate.

As noted by Herring (1950) there is a numerical error in Frenkel's paper which has here been corrected. Since the equations used in the derivation of this formula are linear, the deformation due to a transverse force acting simultaneously on the rod may be considered independently. The foils of Pranatis and Pound may be considered to be unrestrained at the edges. Thus the apparent viscosities of the specimen can be derived from the slopes of the strain rate : stress curves, but the values given by Pranatis and Pound should be divided by three.

It is not clear how Pranatis and Pound have derived their calculated values of the viscosity. On p. 667 of their paper they mis-quote the expression derived by Herring for the apparent viscosity of foils of this type, where the tangential stresses at the grain boundaries are relaxed. The formula should be:

$$\eta = \frac{1}{10}\left(\frac{3}{4\pi}\right)^{2/3}\left(\frac{RT}{D\Omega}\right)V_g^{2/3} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where V_g = average grain volume. Pranatis and Pound do not indicate how they have calculated V_g.

In table 2 the slopes of the strain rate : stress graphs is calculated assuming that

$$V_g = \pi r^2 a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where r is the average radius of the grains and a is the thickness of the foil.

Data on the self-diffusion coefficient of copper as determined by the radioactive tracer technique has been taken from the work of Kuper *et al.* and also from the work of Rollin (1939), so that comparison may be made with the calculations of Pranatis and Pound. 1051°C has been considered as 1050°C, 1024°C as 1025°C, and 1002°C as 1000°C.

Pranatis and Pound give the activation energy for the self-diffusion of copper from their results as 56 800 cal/mole. The best result obtained by the radioactive tracer technique is 47 140 cal/mole.

A number of unweighted single crystal specimens 0·0015 in. thick were heated under conditions where shrinkage would be expected. However, no deformation could be detected.

Table 2

Temperature °c	Foil thickness (in.)	Slope of strain rate : stress graph $cm^2/dyne/sec \times 10^{15}$		
		Observed	Calculated Kuper *et al.*	Calculated Rollin
1050	0·0015	15·7 in He	50·8	22·8
		20·0 in H_2	50·8	22·8
	0·002	10·7	30·9	13·7
	0·003	8·96	22·6	10·0
	0·004	8·96	23·3	10·3
	0·005	4·90	16·8	7·42
1025	0·0015	14·21 in He	36·7	14·8
		14·43 in H_2	36·7	14·8
	0·002	8·26	44·7	8·89
	0·003	7·19	16·4	6·51
1000	0·0015	10·53 in He	26·1	9·44
		10·70 in H_2	26·1	9·44
	0·002	6·21	15·9	5·67
	0·003	4·27	11·4	4·15
975	0·0015	9·01	18·4	5·94
960	0·0015	9·38	14·9	4·44

2.2. *Silver*

Greenough (1952) studied the deformation of silver wires in nitrogen, and found that reproducibility of results was poor. The average strain rate for nine specimens of 0·01016 cm diameter wire heated at 920°c for 741 hours was $-7{\cdot}61 \times 10^{-9}$ per sec, the results for individual specimens ranging from $-9{\cdot}97 \times 10^{-9}$ to $-5{\cdot}02 \times 10^{-9}$ per sec. The predicted strain rate was $-2{\cdot}91 \times 10^{-9}$ per sec, on the basis of the radioactive tracer experiments of Tomizuka and Sonder (1956) on silver single crystals. In a second experiment, the average strain rate in eight specimens at 920°c for 442 hours was $-10{\cdot}6 \times 10^{-9}$ per sec, results ranging from $-13{\cdot}3 \times 10^{-9}$ to $-8{\cdot}28 \times 10^{-9}$ per sec. The predicted strain rate was $-4{\cdot}87 \times 10^{-9}$ per sec.

The deformation of two single crystal wires 0·01 cm diameter with an applied stress of the order of 12×10^4 dynes/cm^2 could not be measured after heating to 910°c for 283 hours. The strain must have been less than

4×10^{-3}. Under the same conditions, the strain in a similar polycrystalline rod having five grains across a diameter was 13×10^{-3}.

Funk *et al.* (1951) studied the deformation of 0·0129 cm diameter silver wires in a helium atmosphere for periods up to 116 hours. From their results, and from the self-diffusion coefficients calculated from the results of Tomizuka and Sonder, the following table has been drawn up.

Table 3

Temperature °C	Slope of strain rate : stress graph cm^2/dyne/sec	
	Observed	Calculated
876 (one result)	$2{\cdot}3 \times 10^{-14}$	$0{\cdot}97 \times 10^{-14}$
907 (three results)	$3{\cdot}7 \times 10^{-14}$	$1{\cdot}57 \times 10^{-14}$
923 (three results)	$6{\cdot}5 \times 10^{-14}$	$1{\cdot}99 \times 10^{-14}$
939 (three results)	$14{\cdot}5 \times 10^{-14}$	$2{\cdot}50 \times 10^{-14}$

The activation energy for the self-diffusion of silver derived from the deformation of the wires is 122 000 cal/mole (omitting the result at 876 °C), or 108 000 cal/mole taking into account all the results.

The work of Buttner *et al.* (1952 a) indicates that a partial pressure of oxygen in the helium of 0·1 atmosphere makes no significant difference to the slope of the strain rate : stress graph at about 935 °C.

2.3. *Gold*

Alexander *et al.* (1951) reported the results of experiments on gold wires, but the experiments were carried out under conditions where the wires must have been rapidly contaminated by nickel and chromium from the nichrome alloys present in the furnace. However, the experimental results are in reasonable agreement with predictions based on Herring's theory. Evidence was presented to show that wires which were shrinking tended to neck and break up into small pieces. The process is illustrated for wires lying on an alundum surface. In this case, friction at the points where the wire touched the surface would give rise to this effect. The apparent bulges and necks observed in sections of suspended wires can be interpreted in terms of the kinks which develop in the wires. No necks would be expected if the wires were shrinking by the Nabbaro–Herring mechanism.

The experimental work of Buttner *et al.* (1952 b) in which 0·0064 cm diameter wire was heated in a helium atmosphere, seems to be more satisfactory. However, their interpretation of their result is unsatisfactory. In their eqn. (6), they derive an expression for the apparent viscosity of the wires by the method of dimensional analysis, but they omit to insert the numerical constant required in this method. The values

which they derive for the apparent viscosity of their wires should thus be divided by three, as reference to the original paper by Udin *et al.* (1949) will show.

In calculating the apparent viscosity from Herring's formula, they ignore the self-diffusion data for gold obtained by the radioactive tracer technique, and reported in a paper by Dienes (1950) which they quote. Instead, they use the activation energy from their own experimental results to calculate the self-diffusion coefficients for gold, on the basis of Dienes' formula (1950).

In table 4 the slopes of the strain rate : stress curves obtained by experiment is compared with the slopes predicted by Herring's formula, using the data obtained by the radioactive tracer technique by Gatos and Kurtz (1954), by Makin *et al.* (1957) and by Okkerse (1956). The last author worked with single crystal specimens.

Table 4

Temperature °C	Slope of strain rate : stress graph $cm^2/dyne/sec \times 10^{14}$			
	Observed	Calculated		
		Gatos *et al.*	Makin *et al.*	Okkerse
1007 (three results)	4·12	2·64	3·74	3·20
1017 (one result)	4·85	3·01	4·22	3·58
1025 (three results)	5·95	3·34	4·64	3·91
1042 (three results)	7·14	4·14	5·63	4·70

The activation energy for deformation obtained by Buttner *et al.* is 51 000 cal/mole. On this basis the activation energy for self-diffusion would be about 55 000 cal/mole. From radioactive tracer experiments the activation energy for self-diffusion has been given as 45 300 cal/mole (Gatos and Kurtz), 41 700 cal/mole (Makin *et al.*) and 39 360 cal/mole (Okkerse), the two latter values probably being the more reliable.

2.4. *Nickel*

Hayward and Greenough (to be published) studied the deformation of 0·01218 cm diameter nickel wire in argon. During the first 30 hours or so, the strain rate in nearly every specimen decreased. Thereafter the rate remained constant to the end of the experiment. The constant rate was generally about eight times faster than the rate calculated from radioactive tracer self-diffusion data, and the self-diffusion activation energy calculated from deformation was 53 000 cal/mole, compared with 66 800 cal/mol from radioactive tracer work (Hoffman *et al.* 1956). A "preliminary rough study" reported by Burgess and Smoluchowski (1955) gives an activation energy in the range 61 000–65 000 cal/mole.

After long periods of deformation, the appearance of the grains suggested that the Nabarro–Herring mechanism was responsible for deformation.

2.5. *Aluminium*

The deformation of 2·54 mm thick aluminium sheet at temperatures in the range 577–647°C has been studied by Harper and Dorn (1957). Unfortunately the experiments were carried out in air. It is probable that the oxide film which grows on the metal surface under these conditions profoundly modifies the processes of generation and absorption of vacant lattice sites on the surface of the metal. It is thus very likely that the deformation process studied here is not strictly comparable with the process considered by Herring, and the process in the other metals here considered. However, it was found that the strain rate decreased with time and then became constant, the constant rate being proportional to the applied stress in the range 3–18 lb/in^2 (2×10^5–13×10^5 dynes/cm^2) approximately.

An interesting feature of this work was the scribing of reference marks on some specimens by a diffraction grating ruling machine. This enabled a detailed study of the deformation to be made. It was shown that the creep curve for specimens so marked agreed very well with that for an unmarked specimen tested under otherwise identical conditions. Grain boundary shearing was shown to account for about 12% of the total deformation. Further, it was argued that if the Nabarro–Herring mechanism was responsible for the deformation, the strain between markers within a grain should be zero, and all the extension of the grains should be concentrated at the transverse grain boundaries.

A specimen was ruled with a series of grid lines 0·003937 in. apart, and a stress of 10 lb/in^2 applied at 647°C until the strain was 0·0170. The average grain diameter was 3·3 mm, so there must have been about 30–40 reference lines on each grain. After deformation, the average strain across transverse grain boundaries was about 0·01695. The implication here is that the distance between grid lines could be measured with an accuracy of at least $\pm 10^{-6}$ in. It is difficult to see how this could be achieved, and the authors give no account of the method of measurement which they used for this purpose. The average strain within the grains was 0·01682.

In a further experiment, the deformation of a single crystal of aluminium at 647°C under an applied stress of 9 lb/in^2 was studied. It appears that the single crystal deformed in much the same way as a polycrystalline specimen under identical conditions, but owing to some mistake, details of comparison have been omitted from the paper.

Comparison of the observed creep rate with that predicted from self-diffusion data is difficult, as no radioactive tracer experiments can be carried out with aluminium. The best estimate for the self-diffusion coefficient of aluminium predicts a slope for the strain rate : stress graph which is about 7×10^{-4} times the observed slope.

The activation energy for deformation is about 35 200 cal/mole. The results give an activation energy for self-diffusion of 36 900 cal/mole.

2.6. *Tin*

Chalmers (1936, 1937) worked in air with cylindrical specimens, 7 cm long, and not less than 3 mm diameter. A central 3 cm length in each specimen was used as the gauge length. Most specimens were single crystals over the gauge length. In other specimens there were a few crystals with longitudinal grain boundaries along the gauge length. Stresses up to 2×10^7 dynes/cm^2 were applied at 21·3°C, and changes in length of the gauge section were measured. At stresses up to 10^5 dynes/cm^2 the strain reached a limiting value of the order of 10^{-5}. The initial rate of creep was proportional to the applied stress, and of the order of 5×10^{-9} per sec for a stress of 10^5 dynes/cm^2. Beyond the limiting strain a very slow deformation rate, of the order of 2×10^{-11} per sec, was observed.

In the specimen with the longitudinal boundaries, no plastic deformation could be detected for stresses below a critical stress dependent on the specimen, but of the order of 10^7 dynes/cm^2.

§ 3. Discussion

In cases where the deformation of the specimens has been observed in detail for periods longer than about 50 hours, it has frequently been found that the rate of deformation decreases during the first 30 hours or so, but thereafter remains constant. This has been observed for silver and nickel wires, and aluminium sheets, but the rate of deformation of copper foils appears to be constant. On the other hand, it is significant that the strains in the smaller diameter copper wires of Udin *et al.* were measured over longer times than for the corresponding larger diameter wires. The strain rates observed in short-time experiments in copper and silver were appreciably higher than those predicted by self-diffusion data, though this is not true for gold.

For copper, gold and nickel there is satisfactory agreement between the values of the activation energy for self-diffusion derived from the deformation results and from radioactive tracer work. There are indications that short-time tests give rather higher values for the activation energy, and this may be the source of the discrepancy for silver.

In some nickel specimens at least, the initially higher strain rate cannot entirely be ascribed to offsetting and kinking at grain boundaries (Herring 1950, Greenough 1952). Extrapolating the linear portion of the strain (ϵ)–time (t) curves to $t = 0$ gives a value for ϵ of about 2×10^{-3}, of which 10^{-3} at the most could be attributed to offsetting and kinking. Another deformation mechanism must operate to give a strain of about 10^{-3} during this initial period.

It is likely that when the stress applied to a specimen is changed, simple edge dislocations act for a time as sources or sinks for vacancies. A simple

estimate shows that this process alone cannot give rise to strains of the required order of magnitude.

When a tensile stress is applied to a crystal, those edge dislocations with their half-planes of atoms perpendicular to the stress can climb by the addition of extra atoms to the half-plane. The dislocation may then be said to act as a source of vacancies, and the crystal extends by the length of the Burgers vector, multiplied by the ratio of the increased area of the half-plane to the area of cross section of the crystal. Edge dislocations on other slip planes and the external surface of the crystal act as sources for the atoms, i.e. sinks for the vacancies. If the climbing edge dislocation is pinned at two nodes, then the dislocation bows out and increases in length (Lomer 1957). Dislocations acting as vacancy sinks must also become curved, but this will be ignored in the following analysis.

The equilibrium radius of curvature of the dislocation line may be determined in the same way as the curvature of a dislocation line in the slip plane when a shear stress is applied. If $\mathbf{b}$ is the Burgers vector, W the energy per unit length and r the radius of curvature of the dislocation, then the applied tensile stress is given by

$$\sigma = \frac{W}{\mathbf{b}r}. \qquad (7)$$

Putting $\sigma = 5 \times 10^5$ dynes/cm^2, $\mathbf{b} = 2 \times 10^{-8}$ cm, $W = 5 \times 10^{-4}$ erg/cm, gives $r = 5 \times 10^{-2}$ cm.

If the average length of edge dislocation between nodes is 10^{-4} cm and the dislocation density 10^8 lines/cm^2, the area of each dislocation 'bow' is about 25×10^{-13} cm^2 and the strain produced in the direction of the tensile stress is at the most 10^{-8}. In order for a dislocation to act as a 'Frank–Read' source or sink for vacancies at this stress (Lomer), the distance between nodes would have to be of the order of 10^{-1} cm. which is greater than the thickness of most of the specimens.

If this mechanism operates at all, the sources must be lengths of edge dislocation rotating about a single node and acting as a 'vacancy mill'. This process would operate most easily if the free end of the dislocation terminated on a free surface of the crystal at every stage in the rotation. A grain boundary could easily prevent the operation of this mechanism, the adjacent grain being unable to accommodate the deformation in the plane of the rotating dislocation. The total deformation would be independent of the magnitude of the applied stress if the rotating dislocation were stopped by structural barriers in the crystal, but the direction of deformation would depend on the direction of the effective applied stress.

On this basis, the higher initial strain rate should be most marked in fine wire specimens with the 'bamboo' structure, and least in foil specimens. The observations of Chalmers on tin can also be explained on this basis.

The work of Barnes *et al.* (1958) on the 'precipitation' of a supersaturated solution of helium in copper pointed to similar conclusions with regard to vacancy sources in the metal.

Acknowledgment

I am grateful to Dr. E. R. Hayward for discussions on this topic, and for details of his experimental work on nickel.

References

Alexander, B. H., Dawson, M. H., and Kling, H. P., 1951, *J. appl. Phys.*, **22,** 439.
Barnes, R. S., Redding, G. B., and Cottrell, A. H., 1958, *Phil. Mag.*, **3,** 97.
Burgess, H., and Smoluchowski, R., 1955, *J. appl. Phys.*, **26,** 491.
Buttner, F. H., Funk, H., and Udin, H., 1952 a, *J. phys. Chem.*, **56,** 657; 1952 b, *Trans. Amer. Inst. min.(metall.) Engrs*, **194,** 401.
Chalmers, B., 1936, *Proc. roy. Soc.* A, **156,** 427; 1937, *J. Inst. Metals*, **61,** 103.
Dienes, G. J., 1950, *J. appl. Phys.*, **21,** 1071.
Frenkel, J., 1945, *J. Phys., Moscow*, **9,** 385.
Funk, E. R., Udin, H., and Wulff, J., 1951, *Trans. Amer. Inst. min. (metall.) Engrs*, **191,** 1206.
Gatos, H. C., and Kurtz, A. D., 1954, *Trans. Amer. Inst. min. (metall.) Engrs*, **200,** 616.
Greenough, A. P., 1952, *Phil. Mag.*, **43,** 1075.
Harpur, J., and Dorn, J. E., 1957, *Acta Met.*, **5,** 654.
Herring, C., 1950, *J. appl. Phys.*, **21,** 459.
Hoffmann, R. E., Pikus, E. W., and Ward, R. A., 1956, *Trans. Amer. Inst. min. (metall.) Engrs*, **206,** 483.
Kuper, A., Letaw, H. Jr., Slifkin, L., Sonder, E., and Tomizuka, C. T., 1954, *Phys. Rev.*, **96,** 1224.
Lomer, W. M., 1957, *Report of a Conference on Vacancies and Other Point Defects in Metals and Alloys* (London: Inst. Metals), p. 94.
Makin, S. M., Rowe, A. H., and Le Claire, A. D., 1957, *Proc. phys. Soc. Lond.* B, **70,** 545.
Nabarro, F. R. N., 1948, *Report of a Conference on the Strength of Solids* (London: Physical Society), p. 75.
Okkerse, B., 1956, *Phys. Rev.*, **103,** 1246.
Pranatis, A. L., and Pound, G. M., 1955, *Trans. Amer. Inst. min. (metall.) Engrs*, **203,** 664.
Rollin, B. V., 1939, *Phys. Rev.*, **55,** 231.
Tomizuka, C. T., and Sonder, E., 1956, *Phys. Rev.*, **103,** 1182.
Udin, H., Shaler, A. J., and Wulff, J., 1949, *Trans. Amer. Inst. min. (metall.) Engrs*, **185,** 186.
Udin, H., 1951, *Trans. Amer. Inst. min. (metall.) Engrs*, **191,** 63.

A New Technique for Decoration of Cleavage and Slip Steps on Ionic Crystal Surfaces†

By G. A. Bassett
Tube Investments Research Laboratories, Hinxton Hall, Cambridge

[Received June 10, 1958]

§ 1. Introduction

The markings and step structures observed on cleavage surfaces of crystals have long been used for a study of crystal imperfections (for example, Zapffe and Worden 1949, Forty 1957). Most observations have been made by optical microscopy or by some form of interferometry. Although the latter method is able accurately to measure the height of steps down to 5 Å in favourable circumstances, steps which are close together on the crystal surface cannot be distinguished (Amelinckx 1951). Surface replicas for study by electron microscopy also do not have sufficient resolution in a vertical direction to detect the smallest steps on crystal surfaces.

In the course of an investigation of the nucleation and growth of epitaxial films a new technique has been discovered which is able to display unit lattice steps on cleavage surfaces of sodium chloride and has the lateral resolution of the electron microscope.

§ 2. Experimental Technique

A small quantity of gold which would give a film of mean thickness 10 Å or less if spread uniformly over the surface is deposited on the crystal surface by vacuum evaporation from a 'V'-shaped tungsten filament. A film of carbon is then deposited (Bradley 1954) on top of the gold layer, the vacuum evaporator being arranged so that the two evaporations can be made without opening the system. The carbon film is removed from the substrate carrying with it the gold deposit from the crystal surface, by lowering gently into water. At the thickness of gold deposit used, a continuous film is not formed but large numbers of small nuclei form on the crystal surface. Nuclei form in greater numbers along the edges of steps on the crystal surface which are thus made visible in the electron microscope as chains of gold particles. On the parts of the crystal surface that are free of steps the nuclei are randomly dispersed.

§ 3. Observations

3.1. *Sodium Chloride*

The pattern of steps revealed by nucleation of the gold on a cleavage surface of rocksalt may vary considerably over even the relatively small

† Communicated by Dr. J. W. Menter. An account of this work was presented at the Conference on the Strength of Whiskers and Thin Films held at Hinxton Hall, and the Cavendish Laboratory, Cambridge, March 17th–19th, 1958.

field of view available in the electron microscope. ($0{\cdot}5 \times 0{\cdot}5$ mm). A typical cleavage step is shown in fig. 1, Pl. 63 decorated with a 5 Å film of gold. The insert in fig. 1 shows part of a cleavage step at a higher magnification. The 'staircase' of which the large step is made up can be readily seen. There are also in the field of view a number of single chains of nuclei; it is considered that these mark single steps on the crystal surface, although not necessarily steps of unit height. Many such single steps are observed following a wandering path eventually to join up with a larger cleavage step on another part of the surface.

It is possible to obtain a direct measure of the height of a cleavage step on the rocksalt surface where the step is crossed by a slip step. In rocksalt slip is generally of the (110) [110] type. The (110) glide plane makes an angle of 45° with the (100) cleavage plane and intersects it along a [001] direction. Slip on the glide plane causes a slip step, the surface trace of which is deflected as it crosses a cleavage step by an amount equal to the height of the cleavage step. Effectively, one may regard the slip step as a marker on the cleavage plane. In fig. 2, Pl. 64, AB can be identified as a slip step since, (*a*) the step terminates on the crystal surface at A where there must be an emergent dislocation having a component of its Burgers vector perpendicular to the cleavage surface, and (*b*) the step is straight for a long distance. As this slip step crosses the cleavage step it is deflected by 150 ± 8 Å. Although the resolution of the decorating particles is not quite adequate for the number of small steps to be determined exactly, it is possible to count 45 ± 2 steps. Now if each of these smaller steps is a monatomic step the height of the cleavage step is 125 Å ± 6 Å (a half unit cell for rocksalt is 2·8 Å). Thus, it may be concluded that most of the individual steps of the 'staircase' making up the large cleavage step are monatomic steps and from this it is considered that steps of atomic height on the crystal surface are decorated by preferential nucleation of the gold particles.

Figure 3, Pl. 64 shows another slip step turning abruptly through 90° and is considered to be an example of cross slip. If this is the correct interpretation of the picture it provides strong evidence for slip on planes different from the usual ones. If slip has occurred on a (110) $[1\bar{1}0]$ system then (001) is the cross slip plane. A number of similar examples have been observed.

Further complex arrays of steps are shown in fig. 4, Pl. 65, which is near a light scratch made on the crystal surface after cleavage. It is clear that if the technique is to be used for a study of plastic deformation more refined methods of deformation must be used.

Rocksalt cleavage surfaces have been examined after deformation by bending at room temperature in air. Decoration replicas from such specimens showed no significant differences from surfaces as cleaved. Slip lines are relatively rare on both types of surface. This observation is consistent with the work of Pratt (1953) who has found that when rocksalt is deformed at room temperature in air, slip appears to be confined below

the surface layers of the crystal and produces a surface rumpling rather than sharp slip steps.

3.2. *Other Materials*

A number of other surfaces have been examined to indicate the range of applicability of this type of decoration technique. Surfaces tried include mica cleavage, lithium fluoride as cleaved and after etching with C.P.4, sodium and potassium chloride whisker surfaces grown from solution and zinc cleavage surface. Although nucleation of the gold deposit has always occurred and it has been possible to remove the gold nuclei adhering to a carbon film, step decoration has been obtained only on lithium fluoride cleavage surfaces. The cleavage of lithium fluoride is smoother than that of rocksalt and a decoration replica of a freshly cleaved lithium fluoride surface shows fewer steps than a rocksalt surface. However after deformation by bending, large numbers of slip lines can be decorated on the tension face (001) of the crystal. These slip lines are all parallel and would be expected from the two sets of (110) slip planes operating. The side face (100) of the crystal shows a 'tartan pattern' of slip steps intersecting at right angles. It is a characteristic of this pattern that the slip steps on one system are sensibly straight whilst the steps on the other system are wavy.

§ 4. Discussion

The mechanism by which the decoration phenomenon occurs is not entirely clear. Gold atoms arriving during vacuum evaporation at the crystal surface have a certain mobility on the surface after arrival. Nucleation apparently occurs more readily along the edges of steps and gives rise to a smaller spacing between nuclei. The size of gold nuclei along step edges is, as might be expected, less than for nuclei formed on flat parts of the crystal substrate. For a gold film with a mean thickness of 3 Å the size of the nuclei along step edges is from 25 Å down to 10 Å; on the flat areas of the crystal nuclei range down in size from about 60 Å. It can be seen that nuclei formed along the terminal step of a cleavage 'staircase', i.e. the step adjoining a flat region of crystal, are intermediate in size between nuclei formed on the flat crystal and in the centre of the cleavage step. This is because the mobility of the gold on the crystal surface permits this terminal step to draw upon mobile material on the flat surface on one side of the step. Some measure of the distance gold atoms can migrate on the crystal surface can be gained from a consideration of the wandering steps on the crystal surface. When two such steps are at a greater distance apart than some critical distance gold nuclei form with a density and size corresponding to the flat regions of the crystal. As the steps get closer together than about 150–200 Å all the arriving gold is able to migrate to nucleating sites along the step edges.

Further work is required to elucidate the manner in which the temperature of the substrate and the rate of deposition influence the nucleation of the deposit layer.

Acknowledgments

I wish to thank my colleagues, Dr. J. W. Menter, Dr. A. J. Forty and Dr. B. H. Kear for helpful discussions in the course of this work. This work is published by permission of the Chairman of Tube Investments Limited.

References

Amelinckx, S., 1951, *Phil. Mag.*, **42,** 324.
Bradley, D. E., 1954, *Brit. J. appl. Phys.*, **5,** 65.
Forty, A. J., 1957, *Proc. roy. Soc.* A, **242,** 392.
Pratt, P. L., 1953, *Discussion Properties of Metallic Surfaces* (London : Institute of Metals, 1954), p. 346.
Zapffe, C. A., and Worden, C. O., 1949, *Acta cryst.*, **2,** 377.

The Knight Shift in Superconductors

By V. Heine and A. B. Pippard†
Royal Society Mond Laboratory, Cambridge

[Received July 15, 1958]

Abstract

The Bardeen, Cooper and Schrieffer (BCS) theory of superconductivity predicts a zero Knight shift at 0°K. An analysis of this result shows how it is related to the exact pairing between electrons of opposite **k** and opposite spin in this theory. The analysis and other considerations suggest a modification of the BCS theory, which then gives a Knight shift in reasonable agreement with Reif's measurements.

Although the experimental evidence concerning the Knight shift in superconducting mercury is not entirely concordant (Reif 1957, Knight *et al.* 1956), the more detailed study of Reif indicates that it is unlikely to vanish at any temperature and that therefore the superconductor at 0°K still retains some of the spin paramagnetism of the normal metal. At first sight this appears to conflict with the theory of superconductivity recently developed by BCS (Bardeen *et al.* 1957 a, b) which predicts zero spin paramagnetism at 0°K (see for instance Yosida 1958). However we wish to show how an analysis of this result suggests that the conflict arises more from the detailed form taken by the BCS theory at present than from any essential weakness. Furthermore, this analysis together with some plausible physical reasoning suggest in an heuristic manner a simple generalization of the BCS theory. On the basis of this generalization we have made a rough calculation of the spin paramagnetism of the superconductor at 0°K and obtained three-quarters of the value for the normal metal, in reasonable agreement with Reif's result.

Let us begin by considering a normal metal containing N conduction electrons in a magnetic field H, and let $\frac{1}{2}N+n$ of the electron spins be directed parallel to the field, so that the moment is $2n\beta$, β being the Bohr magneton. Then the energy of the assembly is

$$E(n)-E(0)=n^2/N(0)-2n\beta H, \qquad (1)$$

where $N(0)$ is the density of states for one spin at the Fermi level; the first term is kinetic energy, the second magnetic potential energy. A minimum of $E(n)$ occurs when $n=N(0)\beta H$, and this gives the usual expression for the Pauli paramagnetism. Now in a superconductor the electrons are in interaction with one another and one must add to (1) a term representing the change in interaction energy consequent upon a reversal of n spins. In the BCS theory there is an exact pairing of electrons

† Communicated by the Authors.

with opposite wave-number and spin, i.e. between the Bloch states $\psi_{\mathbf{k}\uparrow}$ and $\psi_{-\mathbf{k}\downarrow}$, and in the ground state the occupation probabilities $g_{\mathbf{k}\uparrow}$ and $g_{-\mathbf{k}\downarrow}$ are equal, every electron being paired with its opposite number. To turn a spin over involves breaking one pair and this raises the energy of the whole system by a finite amount $2\epsilon_0$ (approximately equal to $3{\cdot}5\,kT_c$ where T_c is the transition temperature of the superconductor). Thus in the strict BCS theory one has to add to (1) an interaction term of the form $2\epsilon_0\,|\,n\,|$. It is now clear that unless $H > \epsilon_0/\beta$ (about 100 kG for mercury), $E(n)$ has a minimum when $n=0$ and the spin paramagnetism vanishes, the spins being locked by the interaction into an antiferromagnetic-like ground-state.

An escape from this situation is provided by any mechanism which replaces the interaction term $2\epsilon_0\,|\,n\,|$ by one which is quadratic in n, and this obviously means that there must be no energy gap of the order of kT_c separating the spin states $n=0$ and $n=1$. This can be achieved by a relaxation of the condition of exact pairing. In qualitative terms this means that when a pair of electrons in the ground state is broken in order to reverse the spin of one, it is not necessary to lose all the interaction energy contributed by this pair; even though they have now no partners of precisely opposite $\mathbf{k}$ and opposite spin, they may benefit nearly as much from interactions with electrons of opposite spin and *nearly* opposite $\mathbf{k}$. This does not appear to be physically unreasonable or in conflict with other aspects of the BCS theory; in particular it is not inconsistent with there being an energy gap of $2\epsilon_0$ for the excitation of an electron across the gap into a 'normal' state ('normal' as opposed to 'superconducting' in the sense of the two-fluid model). We conclude therefore that in order to account for a Knight shift at 0°K, it is necessary to introduce into the theory some correlation between $\mathbf{k}'\uparrow$ and $-\mathbf{k}\downarrow$ for unequal $\mathbf{k}$ and $\mathbf{k}'$, and to count such a correlation as forming at least partially a valid pair.

We can sharpen this conclusion by a more mathematical discussion. Consider the probability $P(\mathbf{k}'\uparrow, -\mathbf{k}\downarrow)$ that the states $\mathbf{k}'\uparrow$ and $-\mathbf{k}\downarrow$ are simultaneously occupied. In the BCS theory in the general case when n may be non-zero, the probability takes the form

$$\begin{aligned} P(\mathbf{k}'\uparrow, -\mathbf{k}\downarrow) &= \langle \Psi_{\mathrm{BCS}}\,|\,c_{\mathbf{k}'\uparrow}{}^{*}c_{\mathbf{k}'\uparrow}c_{-\mathbf{k}\downarrow}{}^{*}c_{-\mathbf{k}\downarrow}\,|\,\Psi_{\mathrm{BCS}}\rangle \\ &= g_{\mathbf{k}'\uparrow}g_{-\mathbf{k}\downarrow} + g_{-\mathbf{k}\downarrow}(1-g_{\mathbf{k}'\uparrow})\delta_{\mathbf{k}\mathbf{k}'}. \qquad \ldots\ldots\ldots \quad (2) \end{aligned}$$

Thus when $\mathbf{k}$ and $\mathbf{k}'$ are not equal the probability is $g_{\mathbf{k}'\uparrow}g_{-\mathbf{k}\downarrow}$, implying no correlation, while when $\mathbf{k}=\mathbf{k}'$ it is $g_{-\mathbf{k}\downarrow}$, implying perfect correlation. Analysing the latter case further, we write

$$\begin{aligned} P(\mathbf{k}\uparrow, -\mathbf{k}\downarrow) &= (g_{-\mathbf{k}\downarrow})(1) \\ &= (\text{probability of } \psi_{-\mathbf{k}\downarrow} \text{ being occupied, i.e. } g_{-\mathbf{k}\downarrow}) \\ &\times (\text{probability that if } \psi_{-\mathbf{k}\downarrow} \text{ is occupied, then } \psi_{\mathbf{k}\uparrow} \text{ is} \\ &\quad \text{also, i.e. unity}). \qquad \ldots\ldots\ldots \quad (3) \end{aligned}$$

Now in the BCS theory, as we saw above, the zero Knight shift is a consequence of the term $2\epsilon_0\,|\,n\,|$ in the energy $E(n)-E(0)$, due to the breaking

of pairs when turning the spins over. In view of the observed Knight shift, we can say therefore that for $n \neq 0$ the lowest energy wave function Ψ_{BCS} in the strict BCS theory contains too few pairs. If we ask how the number of pairs in the wave function can be increased, we see immediately that in (3) the second factor is a probability and cannot be increased beyond its present maximum value of unity. There is *no* chance of introducing greater correlation between $\mathbf{k}\uparrow$ and $-\mathbf{k}\downarrow$ purely by spreading the total spin uniformly over all pairs as in a spin wave, or by any such device. A somewhat similar conclusion has been reached by Yosida (1958). However the amount of correlation (2) can be increased between $-\mathbf{k}$ and $\mathbf{k}'$ ($\neq \mathbf{k}$). We conclude that the explanation of a non-zero Knight shift demands a theory involving correlations between the states $\psi_{\mathbf{k}'\uparrow}$ and $\psi_{-\mathbf{k}\downarrow}$ for $\mathbf{k} \neq \mathbf{k}'$.

We now propose a possible modification of the BCS theory which is based on a more general type of correlation between $\mathbf{k}'\uparrow$ and $-\mathbf{k}\downarrow$ as suggested by the above reasoning, and which gives a finite Knight shift. In (2), we leave alone the first term which is just the statistical uncorrelated probability, and we change the second term. We propose to smooth out the interaction by replacing the $\delta_{\mathbf{k}\mathbf{k}'}$ by a function $F(q)$ of $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ such that $\sum_{\mathbf{q}} F(\mathbf{q}) = 1$. This introduces the required correlation between $-\mathbf{k}\downarrow$ and $\mathbf{k}'\uparrow$, and the success of the original BCS theory suggests that $F(\mathbf{q})$ should be sharply peaked about $\mathbf{q} = 0$. In fact Cantor and Martin (1958) have shown theoretically that $F(q)$ should have a spread in $\mathbf{q}$ of the order of $1/\xi_0$, which is what one would expect physically from the interpretation of ξ_0 as the coherence range. In any case it is difficult to believe that nature prefers a delta function to some such smooth function. It remains to choose the coefficient of $F(\mathbf{q})$. Since the correlation is now spread over a range of energy and of $\mathbf{q}$ which may be taken to be large compared with βH, the simple probability considerations of (3) do not limit us any more. We propose that in a region of size ξ_0 the electron–phonon interaction tries to correlate everything with everything else it can, and therefore that the correlation function is symmetrical between $\mathbf{k}'\uparrow$ and $-\mathbf{k}\downarrow$. This symmetry assumption is sufficient to ensure a Knight shift. It means that the number of pairs is determined not as in (2) by the expression $g_{-\mathbf{k}\downarrow}(1-g_{\mathbf{k}'\uparrow})$ which differs by a term *linear* in n from its value in the ground state, but is determined by some average between $g_{-\mathbf{k}\downarrow}(1-g_{\mathbf{k}'\uparrow})$ and $g_{\mathbf{k}'\uparrow}(1-g_{-\mathbf{k}\downarrow})$. Any average will differ from the ground state ($n=0$) value of $g_{\mathbf{k}}(1-g_{\mathbf{k}})$ only by *second* order quantities in n, and thus lead to an n^2 term in (1) and a non-zero Knight shift. The numerical value of the shift will depend on which particular average is used, and we arbitrarily select the geometric mean. This has the Hartree-like simplicity of being a product of two independent weighting factors, one from $\mathbf{k}$ and one from $\mathbf{k}'$. We propose therefore

$$P(\mathbf{k}'\uparrow, -\mathbf{k}\downarrow) = g_{\mathbf{k}'\uparrow}g_{-\mathbf{k}\downarrow} + [g_{\mathbf{k}'\uparrow}(1-g_{\mathbf{k}'\uparrow})g_{-\mathbf{k}\downarrow}(1-g_{-\mathbf{k}\downarrow})]^{1/2}F(\mathbf{k}-\mathbf{k}') \quad . \; . \; . \quad (4)$$

as a plausible generalization of (2). This expression may be regarded

as the matrix element of $c_{\mathbf{k}'\uparrow}{}^{*}c_{\mathbf{k}'\uparrow}c_{\mathbf{k}\downarrow}{}^{*}c_{-\mathbf{k}\downarrow}$ when the wave function Ψ_{BCS} in (2) is replaced by a supposed 'better' function '|', and by analogy we write down an expression for the slightly more general matrix element $(\mathbf{k}\neq\mathbf{k}')$

$$\begin{aligned}\langle\Psi'|c_{\mathbf{k}\uparrow}{}^{*}c_{-\mathbf{k}+\mathbf{q}\downarrow}{}^{*}c_{-\mathbf{k}'+\mathbf{q}\downarrow}c_{\mathbf{k}'\uparrow}|\Psi''\rangle \\ =[g_{\mathbf{k}\uparrow}(1-g_{\mathbf{k}\uparrow})g_{-\mathbf{k}+\mathbf{q}\downarrow}(1-g_{-\mathbf{k}+\mathbf{q}\downarrow})g_{-\mathbf{k}'+\mathbf{q}\downarrow}(1-g_{-\mathbf{k}'+\mathbf{q}\downarrow}) \\ \times g_{\mathbf{k}'\uparrow}(1-g_{\mathbf{k}'\uparrow})]^{1/4}F(\mathbf{q}). \qquad \ldots\ldots\ldots (5)\end{aligned}$$

It is this matrix element which determines the interaction energy.

We may now estimate the form of $E(n)$ in a superconductor by use of the following assumptions and approximations:

(1) Equation (5) and a constant interaction matrix element V (as in BCS).

(2) $g_{\uparrow}(\epsilon)=g_0(\epsilon-n/N(0)), g_{\downarrow}(\epsilon)=g_0(\epsilon+n/N(0))$, where $g_0(\epsilon)$ is the function (2·35) of BCS (1957 b).

(3) $F(\mathbf{q})$ has a width in energy much less than that of $g_0(\epsilon)$, but H is small enough so that βH is much less than the width of $F(\mathbf{q})$. The first part of this assumption is not justified if the spread of $F(\mathbf{q})$ is of order $1/\xi_0$, since $1/\xi_0$ is comparable with the width of $g_0(\epsilon)$, but the assumption represents a great arithmetical simplification in the evaluation of the integrals.

The calculation is straightforward, and results in replacing $\Gamma(\epsilon)$ by

$$[\Gamma(\epsilon+n/N(0))\Gamma(\epsilon-n/N(0))]^{1/2} \qquad \ldots\ldots (6)$$

in eqn. (4) of BCS (1957 a), which now exhibits explicitly the symmetry between up and down spins. Then on evaluation (1) takes the form

$$E(n)-E(0)=n^2/N(0)-2n\beta H+\tfrac{1}{3}n^2/N(0),$$

the last term being the interaction term. Thus the coefficient of n^2 in (1) is simply multiplied by $\frac{4}{3}$, and correspondingly the spin susceptibility and Knight shift are reduced to three-quarters of their normal value. This is not dissimilar from Reif's value of about two-thirds, and is certainly within the combined limits of error of theory and experiment.

By way of conclusion, we recapitulate briefly the salient features of our discussion. We have assumed the general correctness of the BCS theory of superconductivity but have made the assumption that the interaction between the electrons is completely symmetric between up and down spins, as illustrated by the approximate form (6) of the interaction energy. This assumption, which we suggest is physically plausible, both implies and is implied by a non-zero spin paramagnetism at 0°K and a Knight shift. It means that there is no energy gap for turning a spin over, in contrast to the energy gap for exciting an electron into a 'normal' state. However the assumption is incompatible with the original form of the BCS theory in which correlation is restricted purely to the states $\psi_{\mathbf{k}\uparrow}$ and $\psi_{-\mathbf{k}\downarrow}$ in pairs, because the assumption would imply a probability in (3) being greater than unity: but the incompatibility disappears if the correlation is spread out slightly in $\mathbf{k}$ space. This spreading out is supported also

by independent physical and theoretical arguments, and there is no reason to believe that it affects the existence of an energy gap for excitations to 'normal' states. Thus the characteristic properties derived from electrical conduction (e.g. Meissner effect) may be expected to follow as in the BCS theory except for quantitative corrections. These are probably minor for thermodynamic properties but may be more serious for transport effects, because the latter depend on matrix elements which are more sensitive than energy values to the form of the wave function. We emphasize again that our approach has been entirely heuristic, and in particular we have not established that an actual wave function having our proposed correlation exists. We hope however that this work points out in what direction a theory of superconductivity should be sought, which combines the successes of the BCS theory with an explanation of the Knight shift.

References

Bardeen, J., Cooper, L. N., and Schrieffer, J. R., 1957 a, *Phys. Rev.*, **106**, 162 ; 1957 b, *Ibid.*, **108**, 1175.

Cantor, A. J., and Martin, P. C., 1958, *Bull. Amer. phys. Soc.*, **3**, 202.

Knight, W. D., Androes, G. M., and Hammond, R. H., 1956, *Phys. Rev.*, **104**, 852.

Reif, F., 1957, *Phys. Rev.*, **106**, 208.

Yosida, K., 1958, *Phys. Rev.*, **110**, 769.

CORRESPONDENCE

The Transient Conductivity Increase in Deformed Alkali Halides

By A. Taylor and P. L. Pratt
Department of Physical Metallurgy, University of Birmingham

[Received July 14, 1958]

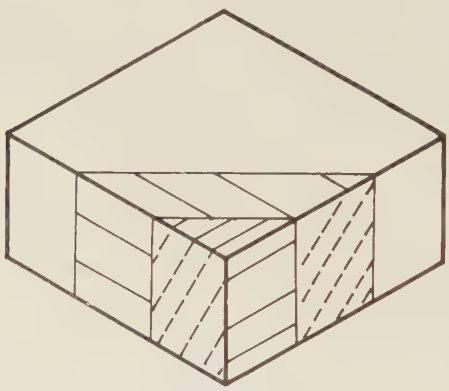
Early observations of the transient effects of plastic deformation upon the ionic conductivity of alkali halide crystals were made by Gyulai and Hartly (1928) and by Stepanow (1933). In these experiments two different effects were observed, a deformation-induced charge-flow without an applied electric field, and a temporary increase in the ionic conductivity. The deformation-induced charge-flow has been studied more recently by Fischbach and Nowick (1955, 1958) and by Caffyn and Goodfellow (1955), and its interpretation in terms of charged dislocations has been discussed elsewhere (Pratt 1957). In this note we wish to comment only on the enhancement of the ionic conductivity produced by deformation.

Seitz (1950, 1952) has interpreted this enhancement of ionic conductivity during compression in terms of the production of lattice vacancies by the intersection of moving dislocations; a deformation of 10% must have generated a density of vacancies of $10^{18}/cm^3$ to account for the 100-fold conductivity increase. Against this Fischbach and Nowick (1958) suggest that the increase is due to the break-up of associated divalent impurity-vacancy complexes, because the number of vacancies appears to them unreasonably large for generation by intersecting dislocations. Further they cite the finding of Caffyn and Goodfellow that the Gyulai–Hartly effect varies widely with the source of the crystal, and is not observed at all in some crystals.

In our experiments on melt-grown crystals, we have found that tall specimens of both NaCl and KCl, typically 1 cm tall by 0·4 cm by 0·4 cm, show no conductivity increases during compression at 90°c up to strains at which the crystals shatter; whereas squat specimens as used by the other workers, typically 0·5 cm square in cross section and 0·2 cm tall, show conductivity increases but only in the later stages of deformation. Following the Russian studies of irrational twinning produced by the compression of similar squat specimens (Brilliantow and Obreimow 1937, Startzev 1940, Classen-Nekludowa 1943), we have examined in some detail the inhomogeneous deformation which takes place. Slip occurs independently on separate (110) $[1\bar{1}0]$ slip systems in different parts of the crystal, in such a way that the crystal is divided macroscopically into blocks. In adjacent blocks slip occurs on systems which meet obliquely.

so that the boundaries between the blocks are (110) planes parallel to the axis of compression, fig. 1. Mixed edge-screw dislocations from the two slip systems meet at the boundary, as Vaughan *et al.* (1958) have also recently observed. Viewed from the side of the compressed specimen, fig. 2, Pl. 66, these boundaries are readily seen. The dislocations retained in the boundary cause a rotation of the lattice on either side about the [111] line of intersection of the two slip systems. This rotation gives rise to tilts on the side of the compressed plate, proportional in magnitude to the total strain, and this led to the Russian description of the effect as 'irrational twinning'. At higher stress levels, slip penetrates the boundaries between the blocks, fig. 3, Pl. 66, and wavy slip lines appear both near the boundary, and within the blocks, fig. 4, Pl. 66. Simultaneously with this penetration of the boundaries, and not before, transient enhancement of the ionic conductivity is observed.

Fig. 1



Formation of blocks by slip on separate systems. Broken lines denote traces of slip planes in sides of specimen.

These results appear to support the interpretation that point defects generated by intersecting dislocations are responsible for the conductivity increase. We hope to publish a fuller account of these experiments shortly.

References

Brilliantow, N. A., and Obreimow, I. W., 1937, *Phys. Z. Sowjet.*, **12**, 7.

Caffyn, J. E., and Goodfellow, T. L., 1955, *Nature, Lond.*, **176**, 878.

Classen-Nekludowa M., 1943, *J. Phys., Moscow*, **7**, 272.

Fischbach, D. B., and Nowick, A. S., 1955, *Phys. Rev.*, **99**, 1333; 1958, *J. phys. chem. Solids*, **5**, 302.

Gyulai, Z., and Hartly, D., 1928, *Z. Phys.*, **51**, 378.

Pratt, P. L., 1957, *Inst. Metals, Monograph and Report Series*, **23**, 99.

Seitz, F., 1950, *Phys. Rev.*, **80**, 239; 1952, *Advanc. Phys.*, **1**, 43.

Startzev, V. T., 1940, *J. Phys., Moscow*, **3**, 107.

Stepanow, A. W., 1933, *Phys. Z. Sowjet.*, **4**, 609.

Vaughan, W. H., Leivo, W. J., and Smoluchowski, R., 1958, *Phys. Rev.*, **110**, 652.

The Nuclear Magnetic Moment of Plutonium-239

By J. Butterworth
Metallurgy Division, A.E.R.E., Harwell, Berks.

[Received May 20, 1958]

It has been shown by optical spectroscopy (Van den Berg and Klinkenberg 1954) and electron paramagnetic resonance methods (Bleaney *et al.* 1954) that the ^{239}Pu nucleus has a spin of $\frac{1}{2}$ and a magnetic moment. By these methods, the best value obtained for the nuclear magnetic moment was $0{\cdot}4 \pm 0{\cdot}2$ nuclear magnetons. To determine the value of the moment more accurately, the nuclear magnetic resonance of ^{239}Pu was sought with a nuclear magnetic induction spectrometer.

Because all the known salts of plutonium are paramagnetic, it seemed most likely that the nuclear magnetic resonance from such salts would be undetectable because of gross broadening. Consequently, the resonance was sought in plutonium metal where it is possible that the paramagnetic electrons are in a conduction band and thus their field at the nuclei is averaged by their motion. Another reason for seeking the resonance in the metal was that the shape of the resonance line could yield information about the unknown electronic structure of the metal.

The sample consisted of just over 9 g of α-plutonium of which about 99% was ^{239}Pu, the rest being mainly ^{240}Pu and ^{241}Pu. Other elements, present as dissolved impurities, included some paramagnetic atoms, but these amounted in number to less than 0·02% of the sample, too small an amount to significantly broaden the resonance. The sample was a powder of particle diameter less than 50 μ to allow the applied radio-frequency field to penetrate it.

The resonance was sought in a direct magnetic field of 4200 gauss at room temperature. To get as large a signal to noise ratio as possible, the dispersion mode of the resonance was sought with the relatively large rotating radio-frequency field of 0·5 gauss, and the bandwidth of the spectrometer was reduced to 1/64 c/s. It was estimated that the signal to noise ratio of the spectrometer was sufficient to allow the resonance of ^{239}Pu to be observed even if the width of the line were as great as 30 gauss.

The search covered the whole range predicted by the electron paramagnetic resonance results, namely 0·2 to 0·6 nuclear magnetons. No resonance was detected.

The absence of the resonance indicates that either the value of the nuclear magnetic moment lies outside the range searched or the resonance line has been broadened out of recognition. Since the present work was completed experimental evidence has been produced that supports both indications. From the results of an atomic beam resonance experiment there has been deduced a value for the nuclear moment of ^{239}Pu of about

0·02 nuclear magnetons (Hubbs *et al.* 1958) while preliminary measurements of the magnetic susceptibility of α-plutonium (Olsen, private communication) appear to give evidence, in the form of a strong negative temperature dependence of the susceptibility and a magnetic transition at a low temperature, for the existence in α-plutonium of localized unpaired electrons. Such electrons would almost certainly cause the obliteration of the nuclear magnetic resonance.

Apart from localized electrons, a number of line-broadening mechanisms are known to be present in α-plutonium, namely, the direct nuclear spin–spin interaction, the nuclear spin-lattice relaxation, the anisotropy of the Knight shift, and the indirect nuclear spin–spin interaction. The extent to which each contributes to the line width cannot be calculated accurately without a more detailed knowledge of the wave functions of the conduction electrons than is available. However, a rough but pessimistic estimate indicates that even their combined effect is not enough to account for the absence of the line. Consequently, the absence of the line, if due to gross broadening, must be attributed to localized unpaired electrons.

The present situation, therefore, is that the value of the nuclear magnetic moment of ^{239}Pu is known with little certainty, and there seems little chance of finding it by nuclear magnetic resonance methods in α-plutonium. The chances of finding the resonance would be greatly increased if a diamagnetic salt of plutonium could be found and used as a sample. Alternatively, if one of the high temperature phases of plutonium were found, the magnetic susceptibility of which was relatively temperature independent, then, provided such a phase could be stabilized at a lower temperature, it would be possible to carry out a nuclear resonance experiment with a better signal to noise ratio and more conveniently than at the higher temperature. If the value of the moment is as low as is indicated by the atomic beam experiment, namely 0·02 nuclear magnetons, then it will be necessary to achieve sample temperatures of the order of 20°K.

Acknowledgments

The author is indebted to Dr. H. M. Finniston for suggesting this work, to him and Mr. A. D. Le Claire for their interest and encouragement, and to Dr. L. E. Drain for helpful discussions. Thanks are also due to Mr. M. J. F. Notley who prepared the sample.

References

Bleaney, B., Llewellyn, P. M., Pryce, M. H. L., and Hall, G. R., 1954, *Phil. Mag.*, **45**, 991.

Hubbs, J. C., Marrus, R., Nierenberg, W. A., and Worcester, J. L., 1958, *Phys. Rev.*, **109**, 390.

Van den Berg, M., and Klinkenberg, P. F. A., 1954, *Physica*, **20**, 461.

REVIEWS OF BOOKS

Progress in Elementary Particle and Cosmic Ray Physics—IV. By J. G. Wilson and S. A. Wouthuysen. (Amsterdam: North-Holland Publishing Company.) [Pp. 470.] 45 guilders.

This volume is the fourth in a series in which the first three were entitled Progress in Cosmic Ray Physics. It contains three articles on Elementary Particles and two on Cosmic Rays. With the development of particle accelerators in the Bev range a new generation of physicists has grown up to whom the properties of new particles and the astrophysical or geophysical aspects of cosmic rays would appear to make strange bedfellows. It seems likely that the greater rate of growth of work on new particles will leave its companion a smaller share in future, which may give further incentive to the issue of separate volumes on the two fields.

The Primary Cosmic Radiation and its Time Variations is described by S. F. Singer in a very readable article which contains some impressive diagrams including a few in the Picasso tradition. The origin of Cosmic Radiation is discussed by V. L. Ginzburg, who describes present theories and makes useful suggestions for further experimental work in both cosmic ray physics and radio astronomy.

Some Theoretical Aspects of the Strong Interactions of the New Particles are described by B. d'Espagnat and J. Prentki, who deal particularly with the possibility of classification by means of three or four dimensional iso-spin space. The "Properties and Production of K-mesons" by W. D. Walker describes experimental work up to about mid-1957, and its interpretation before the non-conservation of parity had become a familiar concept, (though this is just mentioned). The theoretical and experimental aspects of the Interaction of μ-Mesons with Matter are surveyed by G. N. Fowler and A. W. Wolfendale.

R. J. E.

Mechanical Resolution of Linguistic Problems. By A. D. Booth, L. Brandwood and J. P. Cleare. (London: Butterworths Scientific Publications; New York: Academic Press Inc.) [Pp. vii+306.] 50*s.*; $9.80.

This book gives an account of some of the results obtained at Birkbeck College Computational Laboratory on the application of the digital computer APEXC to linguistic problems. A short historical account of mechanical translation from 1947 to the present day is included.

The book proceeds to deal cursorily with such matters as machine methods for making word counts and concordances (apparently the Jesuits are using an I.B.M. computer in the preparation of a concordance to the works of Aquinas) and computer methods for the solution of problems of stylistic analysis (with special reference to the chronological dating of Plato's 'Dialogues'). By far the greatest section of the book (Chapter 9), however, is devoted to what the authors term "a conglomeration of the main problems to be faced in translating German together with one or two suggestions on how to overcome them". This chapter takes, as its basis, the work already done by Oswald and Fletcher on the mechanical resolution of German syntax patterns and endeavours to extend it.

Many of the problems arising from German grammar and syntax are discussed at length and documented with copious examples. Indeed the whole chapter well illustrates the principle that programmers and linguists must work in the closest collaboration in the MT field—never in isolation. No reference is made to detailed programming, this being left for a companion volume.

The final chapter of the book presents an assessment of the special features of a machine built for translation. (A proposed instruction code for such a machine is also given.) It is concluded that the construction of such a machine would require an expenditure of between £50,000 and £100,000.

While the book under review rightly highlights some of the problems involved in MT (more especially in the case of German and, to some lesser extent, French), it cannot be said that the *ad hoc* solutions here proposed for these particular languages will greatly assist in solving the general problems. One is left with the feeling that a more general approach is necessary. M. M.

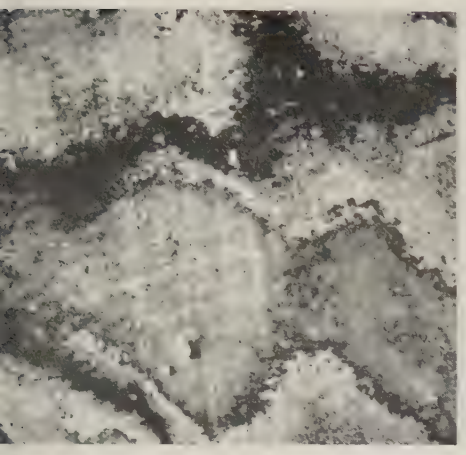
Observation and Interpretation—A Symposium of Philosophers and Physicists. Edited by S. Koerner. (London: Butterworths Scientific Publications.) [Pp. 218.] 40*s*.

This is a discussion of questions belonging to the philosophy of quantum mechanics and borderline problems connected therewith, being a reprint of the papers given together with a verbatim report of the discussions that followed. Two of the seven sections explicitly deal with problems surrounding the conception of probability in its relation to quantum mechanics. One section includes Bohm's attempt at a formulation of quantum theory in terms of hidden variables at a sub-quantum-mechanical level. Relating both this and the former sections is the question whether probability is an ultimate property of a sequence (alternatively, according to Popper, of the experimental arrangement that gives rise to it) or whether the events related by a probability law should be viewed as related by causal laws—at a ' deeper level ' (Bohm). And this point again connects with another, whether a correct theory of measurement should involve a ' realistic interpretation of the formalism of quantum mechanics ' (Feyerabend). In addition to these more technical papers there are others of a more purely ' philosophical ' nature. Often these fit easily into the general scheme; thus the question of the position of the ' observer ', putatively involved in the definition of physical concepts, crops up in the crucial arguments about the foundations of quantum theory (Rosenfeld), the papers on the theory of measurement, but also in those dealing with ' pure ' questions of prediction and determinism.

Participants include Ayer, Bohm, Braithwaite, Popper, Ryle, Rosenfeld, Vigier, Suessmann, Bopp, Groenewold, just to mention some of the 40 odd speakers. The book emphasizes the importance of trying to get clarity on fundamental concepts; above all, it shows physicists and philosophers disputing—thus exhibiting the very nerve of philosophical activity. It makes fascinating reading and should be of equal interest to both physicists and philosophers. G. B.

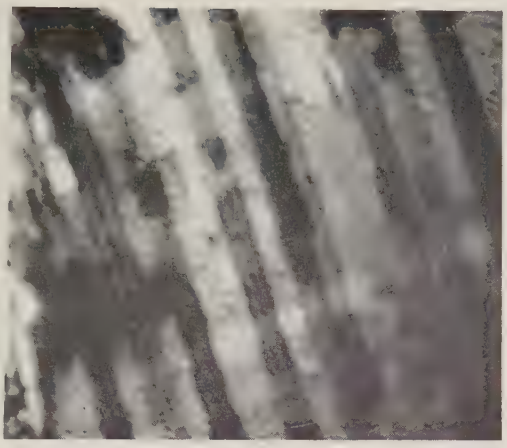
[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

Fig. 3



nted gold film deformed on a single crystal silver substrate. The thinning of the film due to slipping is shown in two directions. ×16 000.

Fig. 4



Dislocation configuration in oriented gold film after deformation on a single crystal silver substrate. Slip on a second system is also evident. ×80 000.

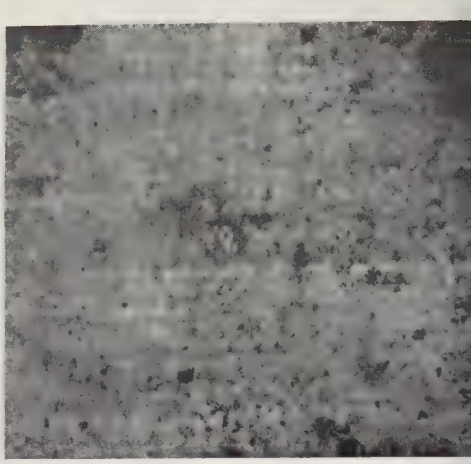
Fig. 5

Slip lines in an oriented gold-palladium alloy film deformed on a single crystal silver substrate. ×16 000.

Fig. 6

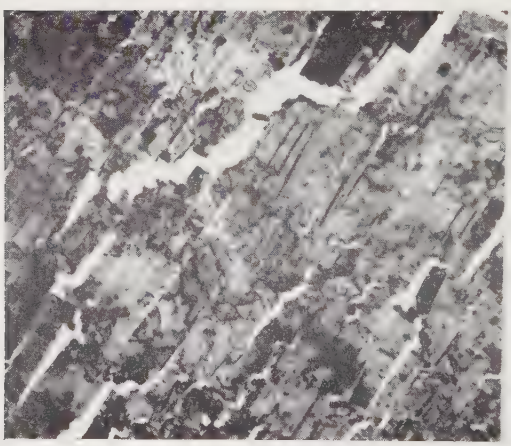
Slip lines and cracking in an oriented platinum film on a single crystal silver substrate. ×16 000.

Fig. 7



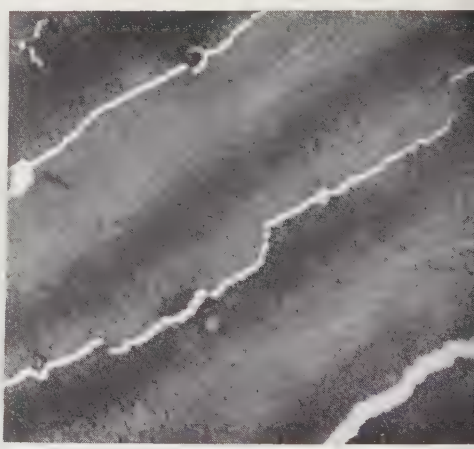
Region of duplex slip in an oriented plati film deformed on a single crystal si substrate. ×16 000.

Fig. 8



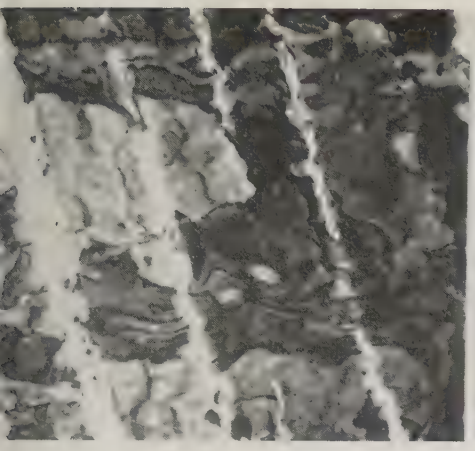
Cracking and slipping of an oriented rhodium film deformed on a single crystal silver substrate. ×16 000.

Fig. 9



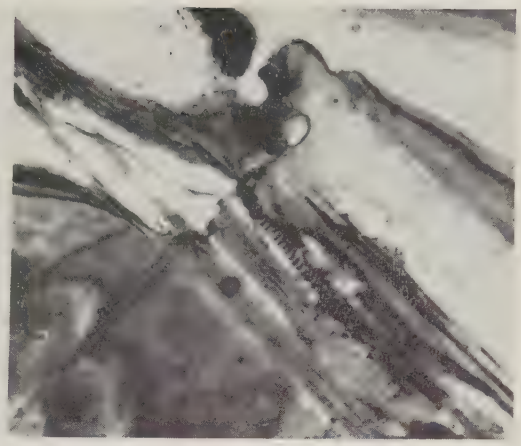
Cracking of oriented rhodium film defor on a single crystal silver substr ×16 000.

Fig. 11



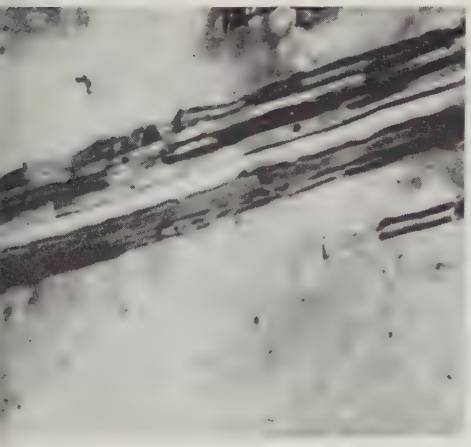
nted gold film which cracked when deformed on a substrate due to an underlying rhodium film acting as a barrier to dislocations. ×16 000.

Fig. 13



Group of dislocations in an oriented gold film deformed on a single crystal silver substrate. A rhodium film acted as a barrier on the other side of the silver substrate. ×80 000.

Fig. 14

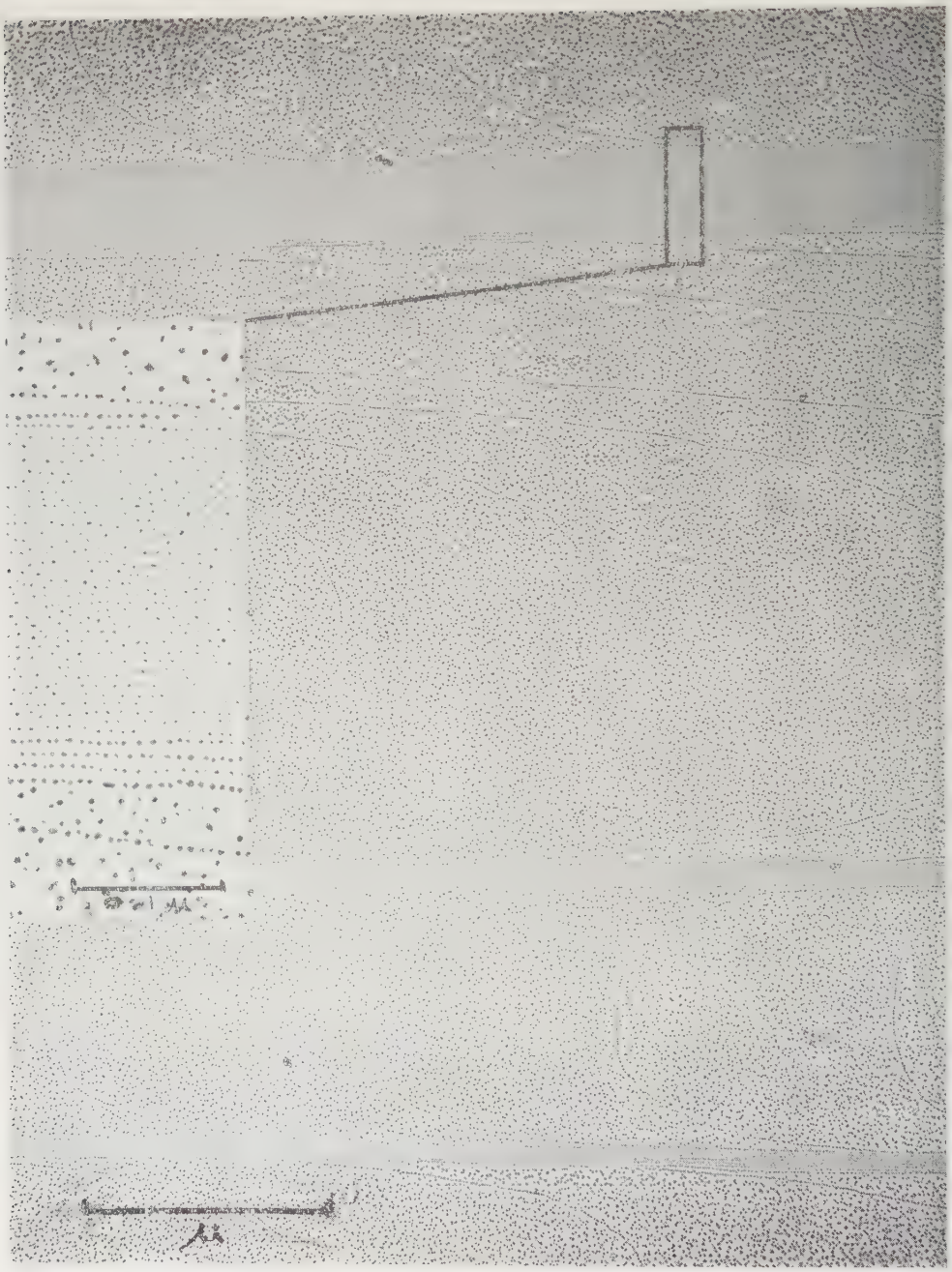


ther group of dislocations in the same specimen as fig. 13. ×80 000.

Fig. 16

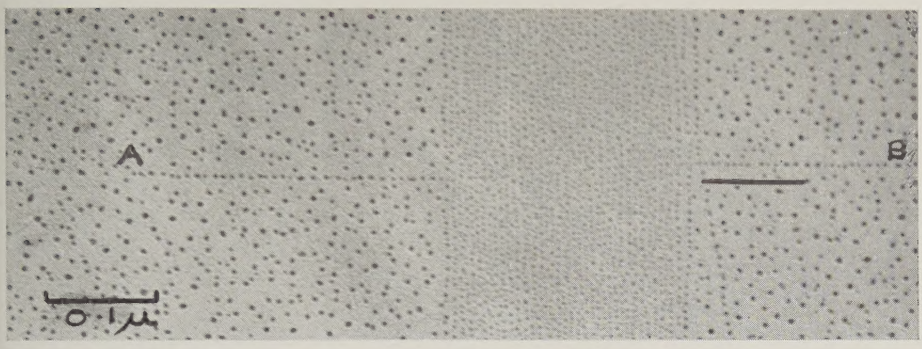
Cracking of polycrystalline rhodium film in one direction. ×16 000.

Fig. 1



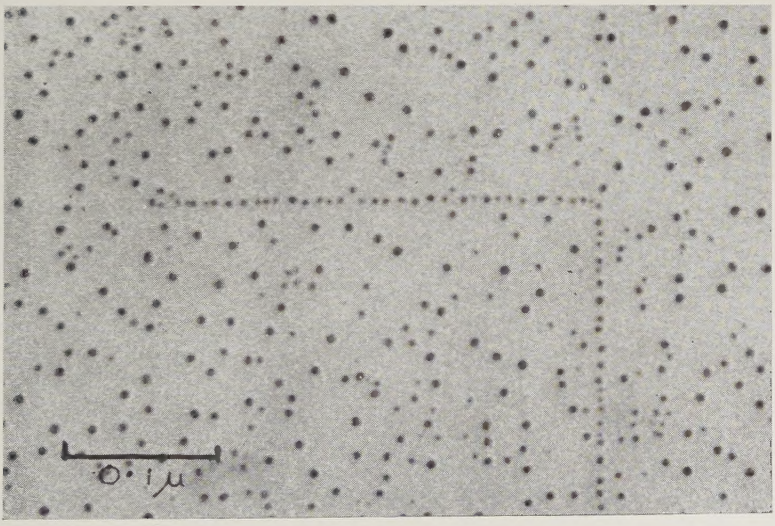
Cleavage steps on cleavage surface of sodium chloride, decorated with gold. $\times 32\,000$. Insert $\times 200\,000$.

Fig. 2



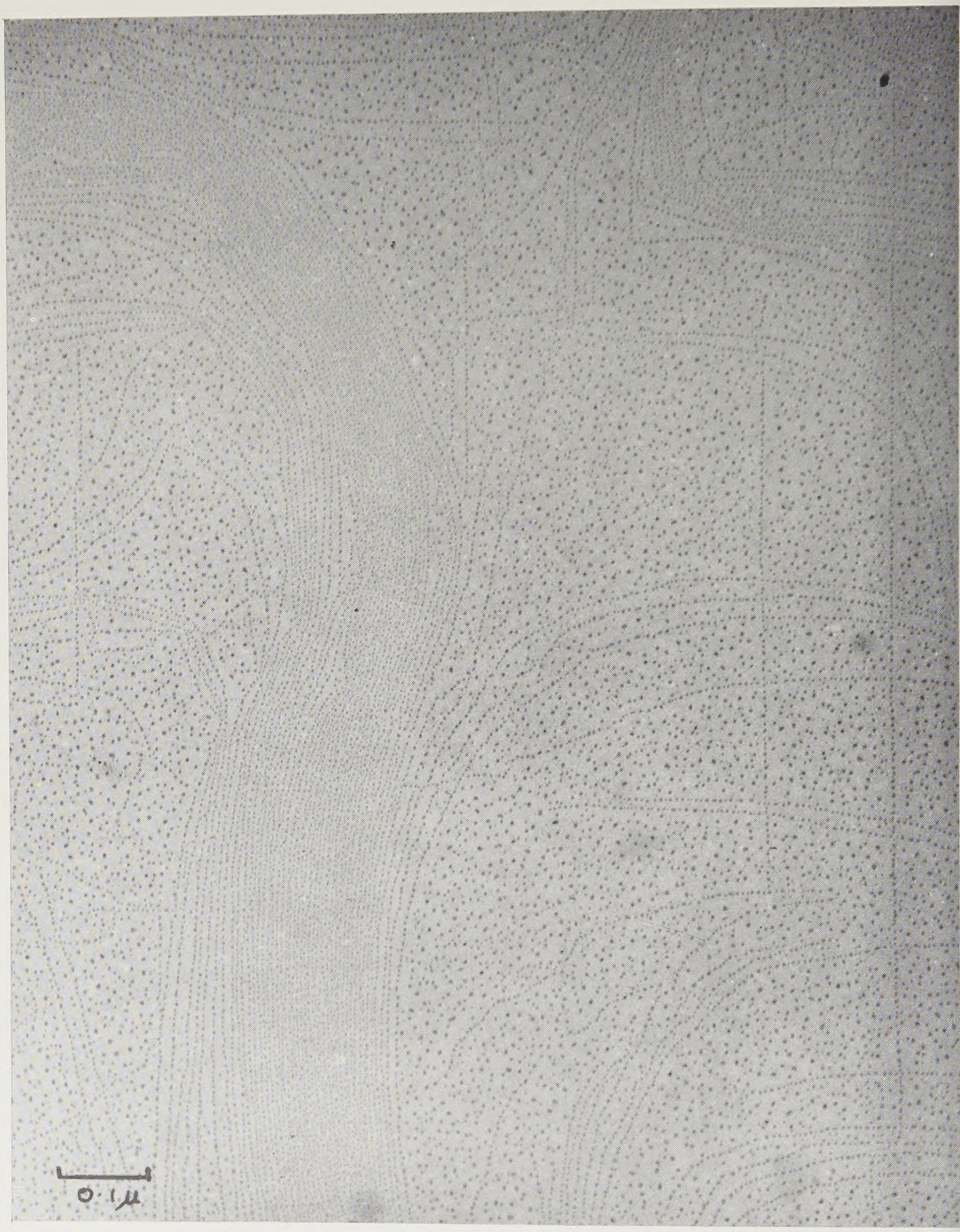
Slip step on sodium chloride deflected on crossing cleavage step. ×140 000.

Fig. 3



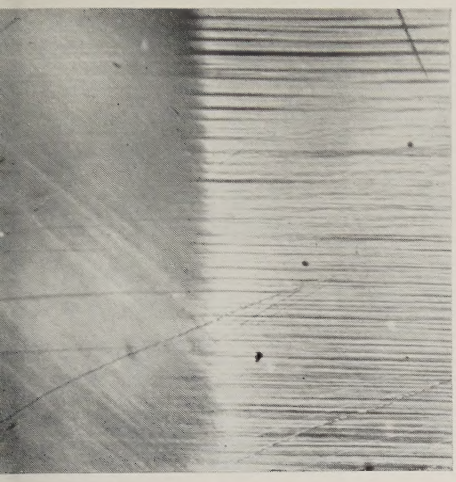
Cross-slip in sodium chloride. ×200 000.

Fig. 4



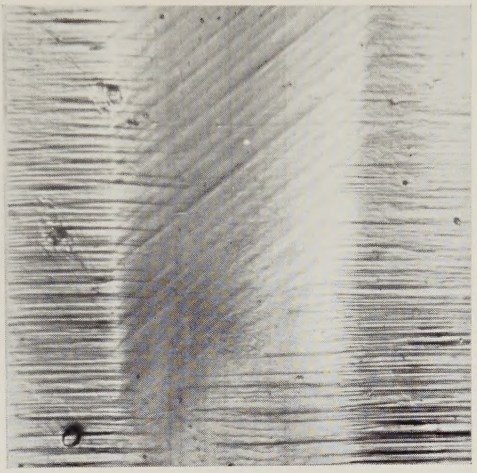
Cleavage and slip steps on sodium chloride near a light scratch on the crystal surface. ×120 000.

Fig. 2



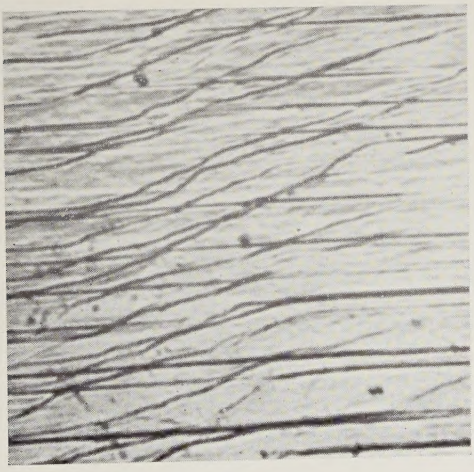
Block boundary before intersection. ×80.

Fig. 3



Block boundary after intersection, showing wavy slip. ×80.

Fig. 4



Wavy slip away from boundary. ×500.